

Appendix A

LENGTH: 352 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
US-09-045-583-52

Query Match 84.8%; Score 952; DB 4; Length 352;
Best Local Similarity 98.9%; Pred. No. 3.9e-85;
Matches 182; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDYQVSSPIYDINVTSEPCQKINVKQIAARLLPPLYSLVIFGFGVGNMVLILINCKR 60
Db 1 MDYQVSSPIYDIDITISPCQKINVKQIAARLLPPLYSLVIFGFGVGNMVLILINCKR 60
QY 61 LKSMTDIYLLNLALISDLFFLLTVPFWAHYAAQWDFGNTMCQLLTGLYFIFGFGSIFII 120
Db 61 LKSMTDIYLLNLALISDLFFLLTVPFWAHYAAQWDFGNTMCQLLTGLYFIFGFGSIFII 120
QY 121 LTTIDRYLVAVHAFALKARTVTGGVTSVITWVAVFASLPGIIFTRSQKGLHYTCSS 180
Db 121 LTTIDRYLVAVHAFALKARTVTGGVTSVITWVAVFASLPGIIFTRSQKGLHYTCSS 180
QY 181 HPPY 184
Db 181 HPPY 184

RESULT 9
US-09-534-185-52
Sequence 52, Application US/09534185
Patent No. 6403767
GENERAL INFORMATION:
APPLICANT: Graham, Gerard J. et al.
TITLE OF INVENTION: No. 6403767el Molecules of the G Protein-Coupled
Heptahelical Receptor Superfamily and Uses
therefor

NUMBER OF SEQUENCES: 56
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD, LLP
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/534,185
FILING DATE: 24-Mar-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/045,583
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Madragouras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: MNI-044
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 52:
SEQUENCE CHARACTERISTICS:
LENGTH: 352 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: Peptide
FRAGMENT TYPE: Internal
SEQUENCE DESCRIPTION: SEQ ID NO: 52:
US-09-534-185-52

Query Match 84.8%; Score 952; DB 4; Length 352;
Best Local Similarity 98.9%; Pred. No. 3.9e-85;
Matches 182; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDYQVSSPIYDINVTSEPCQKINVKQIAARLLPPLYSLVIFGFGVGNMVLILINCKR 60
Db 1 MDYQVSSPIYDIDITISPCQKINVKQIAARLLPPLYSLVIFGFGVGNMVLILINCKR 60
QY 61 LKSMTDIYLLNLALISDLFFLLTVPFWAHYAAQWDFGNTMCQLLTGLYFIFGFGSIFII 120
Db 61 LKSMTDIYLLNLALISDLFFLLTVPFWAHYAAQWDFGNTMCQLLTGLYFIFGFGSIFII 120
QY 121 LTTIDRYLVAVHAFALKARTVTGGVTSVITWVAVFASLPGIIFTRSQKGLHYTCSS 180
Db 121 LTTIDRYLVAVHAFALKARTVTGGVTSVITWVAVFASLPGIIFTRSQKGLHYTCSS 180
QY 181 HPPY 184
Db 181 HPPY 184
RESULT 10
US-08-466-343D-2
Sequence 2, Application US/08466343D
Patent No. 6025154
GENERAL INFORMATION:
APPLICANT: LI, YI
TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING HUMAN G-PROTEIN
CHEMOKINE RECEPTOR HDGNRL0 (AS AMENDED)
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
STREET: 1100 NEW YORK AVE., NW, SUITE 600
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/466,343D
FILING DATE: 06-JUN-1995
CLASSIFICATION: 433
ATTORNEY/AGENT INFORMATION:
NAME: STEFFE, ERIC K.
REGISTRATION NUMBER: 36,688
REFERENCE/DOCKET NUMBER: 1488.1150000/EKS/KLM
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 352 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-466-343D-2

Query Match 84.0%; Score 943; DB 3; Length 352;
Best Local Similarity 97.8%; Pred. No. 2.9e-84;
Matches 180; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 1 MDYQVSSPIYDINVTSEPCQKINVKQIAARLLPPLYSLVIFGFGVGNMVLILINCKR 60
Db 1 MDYQVSSPIYDINVTSEPCQKINVKQIAARLLPPLYSLVIFGFGVGNMVLILINCKR 60
QY 61 LKSMTDIYLLNLALISDLFFLLTVPFWAHYAAQWDFGNTMCQLLTGLYFIFGFGSIFII 120
Db 61 LKSMTDIYLLNLALISDLFFLLTVPFWAHYAAQWDFGNTMCQLLTGLYFIFGFGSIFII 120

QY 121 LTTDRLAVHVAFAKARTVFGVTVITWVAVFASLPGIIFTRSQEGLHYTCSS 180
|||||
Db 121 LTTDRLAVHVAFAKARTVFGVTVITWVAVFASLPGIIFTRSQEGLHYTCSS 180
QY 181 HFPY 184
|||||
Db 181 HFPY 184

RESULT 11

US-09-517-605-5
; Sequence 5, Application US/09517605
; Patent No. 6391567
; GENERAL INFORMATION:
; APPLICANT: Littman, Dan R.
; APPLICANT: Kwon, Douglas S.
; APPLICANT: van Kooyk, Yvette
; APPLICANT: Geijtenbeek, Theo
; TITLE OF INVENTION: METHODS OF USING A FACILITATOR OF RETROVIRAL ENTRY INTO
; TITLE OF INVENTION: CELLS
; FILE REFERENCE: 1049-1-017
; CURRENT APPLICATION NUMBER: US/09/517,605
; CURRENT FILING DATE: 2000-03-02
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 5
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-517-605-5

Query Match 83.4%; Score 936; DB 4; Length 352;

Best Local Similarity 97.3%; Pred. No. 1.4e-83;
Matches 179; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 MDYVSSPIYDINVTSEPCQKINVKQIAARLLPPLYSIVIFGFGVGNMLVILINCKR 60
|||||
Db 1 MDYVSSPTDIDYDSEPCQKINVKQIAARLLPPLYSIVIFGFGVGNMLVILINCKR 60
QY 61 LKSMVDIYLLNLAISDLFFLLVFPFAHAAQWDFGNTMCOLLTGLYFGFGSGLFFII 120
|||||
Db 61 LKSMVDIYLLNLAISDLFFLLVFPFAHAAQWDFGNTMCOLLTGLYFGFGSGLFFII 120
QY 121 LTTDRLAVHVAFAKARTVFGVTVITWVAVFASLPGIIFTRSQEGLHYTCSS 180
|||||
Db 121 LTTDRLAVHVAFAKARTVFGVTVITWVAVFASLPGIIFTRSQEGLHYTCSS 180
QY 181 HFPY 184
|||||
Db 181 HFPY 184

RESULT 12

US-08-724-984A-2
; Sequence 2, Application US/08724984A
; Patent No. 6388055
; GENERAL INFORMATION:
; APPLICANT: Derk Bergsma, Mary Bravner, and Usman Shabon
; TITLE OF INVENTION: No. 6388055el Mouse Genomic Clone of the CC-
; TITLE OF INVENTION: CKR5 Receptor
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road, P.O. Box 1539
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
; COMPUTER: IBM 486
; OPERATING SYSTEM: WINDOWS FOR WORKGROUPS

SOFTWARE: MICROSOFT WORD
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/724,984A
FILING DATE: October 3, 1996
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: William T. Han
REGISTRATION NUMBER: 34,344
REFERENCE/DOCKET NUMBER: ATG50023
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610 270 5024
TELEFAX: 610 270 5090
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 354
TYPE: Amino Acid
TOPOLOGY: Linear
US-08-724-984A-2

Query Match 69.28; Score 776; DB 4; Length 354;

Best Local Similarity 79.0%; Pred. No. 5.3e-68;
Matches 147; Conservative 17; Mismatches 20; Indels 2; Gaps 1;

QY 1 MDYVSSP--YDINVTSEPCQKINVKQIAARLLPPLYSIVIFGFGVGNMLVILINCKR 58
|||||
Db 1 MDYVSSPTDIDYDSEPCQKINVKQIAARLLPPLYSIVIFGFGVGNMLVILINCKR 60
QY 59 KRLKSMVDIYLLNLAISDLFFLLVFPFAHAAQWDFGNTMCOLLTGLYFGFGSGLFFII 118
|||||
Db 61 KRLKSMVDIYLLNLAISDLFFLLVFPFAHAAQWDFGNTMCOLLTGLYFGFGSGLFFII 120
QY 119 ILLTIDRLAVHVAFAKARTVFGVTVITWVAVFASLPGIIFTRSQEGLHYTC 178
|||||
Db 121 ILLTIDRLAVHVAFAKARTVFGVTVITWVAVFASLPGIIFTRSQEGLHYTC 180
QY 179 SSHPY 184
|||||
Db 181 SPPEPH 186

RESULT 13

US-08-466-343D-9
; Sequence 9, Application US/08466343D
; Patent No. 6025154
; GENERAL INFORMATION:
; APPLICANT: LI, Yi
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING HUMAN G-PROTEIN
; TITLE OF INVENTION: CHEMOKINE RECEPTOR HDGMR10 (AS AMENDED)
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVE., NW, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,343D
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: STEFFE, ERIC K.
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1488.1150000/EKS/KLM
; TELECOMMUNICATION INFORMATION:

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 3, 2003, 18:39:24 ; Search time 74.6348 Seconds
(without alignments)
5149.005 Million cell updates/sec

Title: US-09-938-703-3
Perfect score: 2828
Sequence: 1 GAATTCCTCCCAACAGACCA.....AGTAGATTAGATCCGAAATTC 1442

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 1816940

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_n2p.model -DEV=xlp
-O=/cgn2_1/USPTO_spool/US09938703/runat_03062003_161406_22796/app.query.fasta_1.1870
-DB=A_Geneseq 101002 -QFMT=fastan -SUFFIX=rag -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptn -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09938703.qcgn_1_162.runat_03062003_161406_22796 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

A_Geneseq 101002: *
1: /SID52/gcgdata/geneseq/geneseq_emb1/AA1980.DAT: *
2: /SID52/gcgdata/geneseq/geneseq_emb1/AA1981.DAT: *
3: /SID52/gcgdata/geneseq/geneseq_emb1/AA1982.DAT: *
4: /SID52/gcgdata/geneseq/geneseq_emb1/AA1983.DAT: *
5: /SID52/gcgdata/geneseq/geneseq_emb1/AA1984.DAT: *
6: /SID52/gcgdata/geneseq/geneseq_emb1/AA1985.DAT: *
7: /SID52/gcgdata/geneseq/geneseq_emb1/AA1986.DAT: *
8: /SID52/gcgdata/geneseq/geneseq_emb1/AA1987.DAT: *
9: /SID52/gcgdata/geneseq/geneseq_emb1/AA1988.DAT: *
10: /SID52/gcgdata/geneseq/geneseq_emb1/AA1989.DAT: *
11: /SID52/gcgdata/geneseq/geneseq_emb1/AA1990.DAT: *
12: /SID52/gcgdata/geneseq/geneseq_emb1/AA1991.DAT: *
13: /SID52/gcgdata/geneseq/geneseq_emb1/AA1992.DAT: *
14: /SID52/gcgdata/geneseq/geneseq_emb1/AA1993.DAT: *
15: /SID52/gcgdata/geneseq/geneseq_emb1/AA1994.DAT: *
16: /SID52/gcgdata/geneseq/geneseq_emb1/AA1995.DAT: *
17: /SID52/gcgdata/geneseq/geneseq_emb1/AA1996.DAT: *
18: /SID52/gcgdata/geneseq/geneseq_emb1/AA1997.DAT: *
19: /SID52/gcgdata/geneseq/geneseq_emb1/AA1998.DAT: *
20: /SID52/gcgdata/geneseq/geneseq_emb1/AA1999.DAT: *
21: /SID52/gcgdata/geneseq/geneseq_emb1/AA2000.DAT: *
22: /SID52/gcgdata/geneseq/geneseq_emb1/AA2001.DAT: *
23: /SID52/gcgdata/geneseq/geneseq_emb1/AA2002.DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------------|---------------------|
| 1 | 1768.5 | 67.3 | 371 | 19 AAW23834 | Human CC chemokine |
| 2 | 1745.5 | 66.4 | 352 | 18 AAW27407 | Human CCR5. Homo |
| 3 | 1745.5 | 66.4 | 352 | 18 AAW27123 | Human chemokine re |
| 4 | 1745.5 | 66.4 | 352 | 19 AAW23835 | Human CC chemokine |
| 5 | 1745.5 | 66.4 | 352 | 20 AAW88232 | HIV-1 co-receptor |
| 6 | 1745.5 | 66.4 | 352 | 22 AAG80111 | Human CCR5 protein |
| 7 | 1745.5 | 66.4 | 352 | 22 AAB82948 | Human HIV-1 co-rec |
| 8 | 1745.5 | 66.4 | 352 | 22 AAB83354 | Human CCR5 protein |
| 9 | 1745.5 | 66.4 | 352 | 22 AAE04321 | Human chemokine (C |
| 10 | 1745.5 | 66.4 | 352 | 23 ABB08343 | Human CC chemokine |
| 11 | 1745.5 | 66.4 | 352 | 23 AAM52828 | Fusion protein con |
| 12 | 1745.5 | 66.4 | 439 | 20 AAY41280 | Non-endogenous hum |
| 13 | 1739.5 | 66.2 | 352 | 22 ABB56342 | Human CCR5 Gln 55 |
| 14 | 1739.5 | 66.2 | 352 | 23 AAM52829 | Human G-protein ch |
| 15 | 1737.5 | 66.1 | 352 | 22 AAE07039 | Human G-protein ch |
| 16 | 1737.5 | 66.1 | 352 | 22 AAE07048 | Human HDGRI10 prot |
| 17 | 1737.5 | 66.1 | 352 | 22 AAB46858 | Human G-protein ch |
| 18 | 1737.5 | 66.1 | 352 | 23 AAG97152 | Human G-protein ch |
| 19 | 1730.5 | 65.8 | 352 | 18 AAW07602 | Human G-protein ch |
| 20 | 1730.5 | 65.8 | 352 | 21 AAY80128 | Human G-protein ch |
| 21 | 1730.5 | 65.8 | 352 | 22 AAE07037 | Human G-protein ch |
| 22 | 1730.5 | 65.8 | 352 | 22 AAE07046 | Human G-protein ch |
| 23 | 1730.5 | 65.8 | 352 | 23 AAG97150 | Human G-protein ch |
| 24 | 1718.5 | 65.4 | 352 | 22 AAG75089 | Amino acid sequenc |
| 25 | 1711.5 | 65.1 | 352 | 18 AAW27125 | Macaque chemokine |
| 26 | 1670 | 63.5 | 332 | 18 AAW26766 | Human chemokine re |
| 27 | 1463.5 | 55.7 | 354 | 19 AAW54037 | Mouse CC-CKR5 prot |
| 28 | 1332 | 50.7 | 360 | 22 AAG07613 | Human CCR2-641 pol |
| 29 | 1331 | 50.6 | 360 | 16 AAR79166 | Human monocyte che |
| 30 | 1331 | 50.6 | 360 | 18 AAW35833 | Human monocyte che |
| 31 | 1331 | 50.6 | 360 | 22 AAG80108 | Human CCR2b protei |
| 32 | 1331 | 50.6 | 360 | 22 AAG07614 | Human wild-type CC |
| 33 | 1325 | 50.4 | 360 | 22 ABB56340 | Non-endogenous hum |
| 34 | 1191 | 45.3 | 374 | 16 AAR79165 | Human monocyte che |
| 35 | 1191 | 45.3 | 374 | 22 AAG80107 | Human CCR2a protei |
| 36 | 1122 | 42.7 | 215 | 18 AAW27408 | Inactive human CCR |
| 37 | 1122 | 42.7 | 215 | 20 AAW88238 | HIV-1 co-receptor |
| 38 | 1105 | 42.0 | 329 | 22 AAB46859 | Human MCP-1 recept |
| 39 | 1003.5 | 38.2 | 355 | 15 AAF52749 | C-C chemokine rece |
| 40 | 1003.5 | 38.2 | 355 | 18 AAW26588 | Human MIP-1 alpha/R |
| 41 | 1003.5 | 38.2 | 355 | 18 AAW25751 | Human MIP-1 alpha/R |
| 42 | 1003.5 | 38.2 | 355 | 21 AAB20571 | Human CC chemokine |
| 43 | 965 | 36.7 | 355 | 18 AAW29179 | Rat CC chemokine r |
| 44 | 958 | 36.5 | 184 | 18 AAW27406 | Inactive human CCR |
| 45 | 916 | 34.9 | 355 | 19 AAW51744 | Human C-C chemokin |

ALIGNMENTS

RESULT 1
AAW23834
ID AAW23834 standard; Protein; 371 AA.
XX
AC AAW23834;
XX
DT 08-JUN-1998 (first entry)
XX
DE Human CC chemokine receptor 5 (CCR5) A127V variant.
XX
KW CC chemokine receptor 5; CCR5; G-protein coupled receptor;
KW human immunodeficiency virus; HIV; CCR5; AIDS; therapy;
KW transgenic animal.
XX
OS Homo sapiens.
XX
EH Key Location/Qualifiers
FT Domain 48..74
FT /label= I

QY 288 TCGGAGCCCTGCCAAAATCAATGTGAAGCAATCGAGCCCGCCTCCTGCCCTCCGCTC 347

XX

XX DE Human CCR5.
 XX DE
 KW Human Cys-Cys chemokine receptor 5; CCR5;
 KW human immunodeficiency virus; type 1; type 2; HIV-1; HIV-2;
 KW diagnosis; treatment; prevention;
 KW inflammatory disease; rheumatoid arthritis; glomerulonephritis;
 KW asthma; idiopathic pulmonary fibrosis; psoriasis; viral infection;
 KW cancer; atherosclerosis; autoimmune disorder.
 XX OS
 XX OS Homo sapiens.
 XX PN W09732019-A2.
 XX PD
 XX PD 04-SEP-1997.
 XX PF 28-FEB-1997; 97WO-BE00023.
 XX PR 06-AUG-1996; 96EP-0870102.
 XX PR 01-MAR-1996; 96EP-0870021.
 XX PR (EURO-) EUROSREEN SA.
 XX PA
 XX PI Libert F, Parmentier M, Samson M, Vassart G;
 XX PI
 XX PR WPI; 1997-479829/44.
 XX DR N-PSDB; AAT90117.
 XX PT Active and inactive forms of human CC chemokine receptor CCR-5 -
 XX PT useful to diagnose, prevent and/or treat inflammatory disorders,
 XX PT autoimmune disease and viral infection
 XX PS Claim 4; Fig 1b-c; 94pp; English.
 XX CC
 XX CC The present sequence is human CC (Cys-Cys) chemokine receptor
 CC 5 (CCR5), which is stimulated by MIP-1 alpha, MIP-1 beta or RANTES
 CC chemokines, but not by monocyte chemoattractant protein 1 (MCP-1),
 CC MCP-2, MCP-3, interleukin-8 (IL-8) or growth related gene product
 CC alpha (GRO alpha) chemokines. Active CCR-5 is also a receptor of
 CC human immunodeficiency virus type 1 or type 2 (HIV-1 or HIV-2).
 CC CCR5 or its cDNA can be used to diagnose, treat and/or prevent
 CC inflammatory diseases, e.g. rheumatoid arthritis,
 CC glomerulonephritis, asthma, idiopathic pulmonary fibrosis and
 CC psoriasis, viral infections, especially HIV-1 or HIV-2 infection,
 CC cancer, atherosclerosis and autoimmune disorders.
 XX SQ Sequence 352 AA;
 Alignment Scores:
 Pred. No.: 3,38e-187 Length: 352
 Score: 1745.50 Matches: 340
 Percent Similarity: 96.60% Conservative: 1
 Best Local Similarity: 96.32% Mismatches: 0
 Query Match: 66.42% Indels: 12
 DB: 18 Gaps: 1
 US-09-938-703-3 (1-1442) x AAW27407 (1-352)
 QY 240 ATGATATACAGTGTCAAGTCCATCTATGACATCAATATTATATACATCGAGCCCTGC 299
 DB 1 MetAspTyrGlnValSerSerProIleTyrAspIleAsnTyrTyrThrSerGluProCys 20
 QY 300 CAAAATATCAATGTGAGCAATATCGAGCCGCTCCCTGCTCGCTCACTCACTGGTG 359
 DB 21 GlnLysIleAsnValIysGlnIleAlaIleArgLeuLeuProLeuTyrSerLeuVal 40
 QY 360 TTCATCTTTGGTTTGGCAATCTGCTGCTCATCTCACTCACTGATAACTGCAAAAGG 419
 DB 41 PheIlePheGlyPheValGlyAsnMetLeuValIleLeuIleLeuIleAsnCysLysArg 60
 QY 420 CTGAAGAGCATGATGACATCTACCTGCTCAACCTGGCCATCTCTGACCTGTTTCCTT 479
 DB 61 LeuLysSerMetThrAspIleTyrLeuLeuAsnLeuAlaIleSerAspLeuPheLeu 80

QY 480 CTACTGTCCTCCCTTCTGGCTCACTATGCTGCCGCCAGTGGGACTTTTGGAAATACAATG 539
 DB 81 LeuThrValProPheTrpAlaHisTyrAlaAlaGlnTrpAspPheGlyAsnThrMet 100
 QY 540 TGCTCACTCTTCCAGAGGCTCTATTTTATAGGCTTCTCTCTGGAATCTTCTCATCATC 599
 DB 101 CysGlnLeuLeuThrGlyLeuTyrPheIleGlyPhePheSerGlyIlePhePheIleIle 120
 QY 600 CTCTGACAAATCGATAGTACCTGGCTGCTGCATGCTGTGTTGCTTTAAAGCCAGG 659
 DB 121 LeuLeuThrIleAspArgTyrLeuAlaValValHisAlaValPheAlaLeuLysAlaArg 140
 QY 660 ACGGTCACTTGGGCTGCTGACAAAGTGTGATCACTGGGTGGTGGCTGTGTCCTCT 719
 DB 141 ThrValThrPheGlyValValThrSerValIleThrTrpValValAlaValPheAlaSer 160
 QY 720 CTCCAGGAATCATCTTTACCATGCTCAAAAGAGAGCTCTTCAATACACCTGACGCTCT 779
 DB 161 LeuProGlyIleIlePheThrArgSerGlnLysGlnLysGlnLysTyrThrCysSerSer 180
 QY 780 CATTTTCCATPAC-----ATTAAGATAGTCAT 806
 DB 181 HisPheProTyrSerGlnTyrGlnPheTrpLysAsnPheGlnThrLeu-LysIleValIle 200
 QY 807 CTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 866
 DB 200 eLeuGlyLeuValLeuProLeuLeuValMetValIleCysTyrSerGlyIleLeuLysTh 220
 QY 867 TGTGCTTGGTGTGCGAAATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 926
 DB 220 rLeuLeuArgCysArgAsnGluLysLysArgHisArgAlaValArgLeuIlePheThrIle 240
 QY 927 CATGATTTGTTTATTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 986
 DB 240 eMetIleValTyrPheLeuPheTrpAlaProTyrAsnIleValLeuLeuLeuAsnThrPh 260
 QY 987 CCAGGAATTTCTTGGCTGTAATAATTGCAGTAGCTCTAACAGTTGGACCAAGCTATGCA 1046
 DB 260 eGlnGluPhePheGlyLeuAsnAsnCysSerSerSerAsnArgLeuAspGlnAlaMetGl 280
 QY 1047 GGTGACAGACACTTGGGATGACGCACTGCTGCTGATCAACCCCATCATCTATGCTTTGT 1106
 DB 280 nValThrGluThrLeuGlyMetThrHisCysCysIleAsnProIleIleTyrAlaPheVa 300
 QY 1107 CGGGAGAGACTTCAGAACTACCTCTTAGTCTTCTCCAAAAGCACATTTGCCAACGCTT 1166
 DB 300 lGlyGluLysPheArgAsnTyrLeuLeuValPheGlnLysHisIleAlaLysArgPh 320
 QY 1167 CTGCAATGCTGTCTTATTTTCCAGCAAGAGGCTCCGAGCGAGCAAGCTCAAGTTTACAC 1226
 DB 320 eCysLysCysCysSerIlePheGlnGlnGluAlaProGluArgAlaSerSerValTyrTh 340
 QY 1227 CCGATCCACTGGGAGGAGCAAGAAATATCTGTGGGCTTG 1263
 DB 340 rArgSerThrGlyGluGlnIleSerValGlyLeu 352
 RESULT 3
 AAW27123
 ID AAW27123 standard; Protein; 352 AA.
 XX AC
 XX AAW27123;
 DT 14-DEC-1997 (first entry)
 XX DE Human chemokine receptor 88C.
 XX KW Chemokine receptor 88C; atherosclerosis; rheumatoid arthritis;
 KW tumour; asthma; viral infection; AIDS; inflammation;
 KW autoimmune disease; therapy; diagnosis; leukocyte trafficking;
 KW G protein coupled receptor; ligand; modulator; antibody; human.
 XX OS Homo sapiens.

| | | |
|---|--|------------------------------|
| XX | Key | Location/Qualifiers |
| PH | Domain | 1..32 |
| FT | /label= Extracellular_domain | 56..67 |
| FT | Domain | /label= Intracellular_domain |
| FT | Domain | 89..112 |
| FT | Domain | /label= Extracellular_domain |
| FT | Domain | 125..145 |
| FT | Domain | /label= Intracellular_domain |
| FT | Domain | 166..191 |
| FT | Domain | /label= Extracellular_domain |
| FT | Domain | 213..235 |
| FT | Domain | /label= Intracellular_domain |
| FT | Domain | 259..280 |
| FT | Domain | /label= Extracellular_domain |
| FT | Domain | 301..352 |
| FT | Domain | /label= Intracellular_domain |
| XX | W09722698-A2. | |
| PN | 26-JUN-1997. | |
| XX | 20-DEC-1996; | 96WO-US20759. |
| XX | 07-JUN-1996; | 96US-0661393. |
| PR | 20-DEC-1995; | 95US-0575967. |
| XX | (ICOS-) ICOS CORP. | |
| XX | Gray PW, Raport CJ, Schweickart VL; | |
| FI | WPI; 1997-341689/31. | |
| DR | N-PSDB; AAT85161. | |
| DR | New nucleic acid encoding chemokine receptors 88-2B and 88C - used to modulate leukocyte trafficking, e.g. for treatment of inflammation, tumours, viral infections, autoimmune diseases, etc. | |
| XX | ClaIm 16; Page 47-48; 65pp; English. | |
| XX | This polypeptide sequence comprises novel human chemokine receptor 88C, a G protein coupled receptor that is involved in leukocyte trafficking. Its amino sequence was deduced from a cDNA clone (AAT85161) isolated from a macrophage library. It shows 62% identity to CCXCR1. Chemokine receptor 88-2B (see AAW27124) has also been identified. 88C and 88-2B receptors and their polypeptide fragments can be produced in transformed host cells. The receptors, peptides comprising one or more of the extracellular or intracellular domains, and anti-receptor antibodies can be used to modulate receptor activities, particularly ligand and G protein binding, and are potentially useful in the treatment of atherosclerosis, rheumatoid arthritis, tumours, asthma, viral infection, AIDS, inflammatory conditions, pathological immune response, abnormal haematopoietic processes etc. | |
| XX | Sequence | 352 AA; |
| SQ | Alignment Scores: | |
| | Pred. No.: | 3.38e-187 Length: 352 |
| | Score: | 1745.50 Matches: 340 |
| | Percent Similarity: | 96.60% Conservative: 1 |
| | Best Local Similarity: | 96.32% Mismatches: 0 |
| | Query Match: | 56.42% Indels: 12 |
| | DB: | 18 Gaps: 1 |
| US-09-938-703-3 (1-1442) x AAW27123 (1-352) | | |
| QY | 240 ATGGATTATCAAGTGTCACGTCCATCTATGACATCAATATTATATCATCGGAGGCCCTGC | 299 |
| Dd | | |
| Dd | 1 MetAspTyrGlnValSerSerProIleIleArgPheLeuHisThrSerGluProCys | 20 |
| QY | 300 CAAAAATCAATGTGAAGCAAAATCGACGGCGGCTCGCTCGCTCTACTACTGTTG | 359 |

Db 300 lGlyGlyPheArgAsnTyrLeuValPheGlnLysHisIleAlaLysArgPh 320
 QY 1167 CTGAAATCTTCTTATTTCCAGCAGAGGCTCCGAGCGAGCAAGCTCAGTTTACAC 1226
 Db 320 eCysLysCysSerIlePheGlnGlnAlaProGluArgAlaSerValTyrTh 340
 QY 1227 CCGATCCATCTGGGAGCAGCAAAATATCTGTGGGCTTG 1263
 Db 340 rArgSerThrGlyGlnGlnGluLeuSerValGlyLeu 352

RESULT 5

AAW88232
 ID AAW88232 standard; Protein; 352 AA.

AC AAW88232;

DT 15-MAR-1999 (first entry)

XX HIV-1 co-receptor CCR5.

KW HIV-1; CCR5; CCR5m303; co-receptor; infection; diagnosis; AIDS;
 KW gene therapy; human.

OS Homo sapiens.

EH Location/Qualifiers

FT Domain 32..56

FT /note= "transmembrane domain 1"

FT Domain 67..87

FT /note= "transmembrane domain 2"

FT Misc-difference 101

FT /note= "corresponds to TGT (Cys) in wild-type CCR5,

FT TGA (Stop) in CCR5m303"

FT Domain 103..124

FT /note= "transmembrane domain 3"

FT Domain 142..167

FT /note= "transmembrane domain 4"

FT Domain 200..223

FT /note= "transmembrane domain 5"

FT Domain 236..260

FT /note= "transmembrane domain 6"

FT Domain 275..301

FT /note= "transmembrane domain 7"

FW W09854317-Al.

PD 03-DEC-1998.

XX 29-MAY-1998; 98WO-EP03437.

XX 30-MAY-1997; 97US-0048057.

XX (MOND-) FOND MONDIALE RECH & PREVENTION SIDA.

PA Arenzana Siesdedos F, Beretta A, Braun J, Quillent C;

XX WPI; 1999-059835/05.

DR N-PSDB; AAV84126.

XX New CCR5 variant protein of the HIV-1 co-receptor - useful in

PT developing resistance of CCR5-expressing cells to HIV-1 infection

XX Disclosure; Page 34-35; 55pp; English.

CC strains, and may indicate slower progression of the disease. The
 CC detection of CCR5 variants may be used to identify individuals at
 CC lower risk of infection relative to the general population who, if
 CC infected, may exhibit slower progression to AIDS. Probes and
 CC primers (see AAV84127-36) are provided for use in diagnostic methods
 CC for detecting the presence of such variants. A method is provided
 CC for inhibiting HIV-1 infection of a cell expressing the CCR5
 CC receptor. This involves introducing a nucleic acid encoding a CCR5
 CC variant into the cell, thereby reducing the number of functional
 CC CCR5 molecules present on the cell surface.

XX Sequence 352 AA;

Alignment Scores:

Pred. No.: 3-38e-187 Length: 352
 Score: 1745.50 Matches: 340
 Percent Similarity: 96.60% Conservative: 1
 Best Local Similarity: 96.32% Mismatches: 0
 Query Match: 66.42% Indels: 12
 DB: 20 Gaps: 1

US-09-938-703-3 (1-1442) x AAW88232 (1-352)

QY 240 ATGGATTATCAAGTCTCAAGTCCCAATCTATGACATCAATATTATATACATCGAGCCCTGC 299
 Db 1 MetAspTyrGlnValSerSerProIleTyrAspIleAsnTyrTyrThrSerGluProCys 20
 QY 300 CAAAAATCAATGGAAGCAATCGCAGCCGCTCCTCGCTCGCTACTACTACTGCTG 359
 Db 21 GlnLysIleAsnValLysGlnIleAlaAargLeuLeuProLeuTyrSerLeuVal 40
 QY 360 TTCATCTTTGTTTGTGGGCAACATGCTGGTCACTCCCTCATCTGATAAATGCGAAAGG 419
 Db 41 PheIlePheGlyPheValGlyAsnMetLeuValIleLeuIleLeuLeuLeuLeuLeuLeu 60
 QY 420 CTGAAGAGCATGACTGACATCTACCTGCTCAACTGGCATCTCTGACCTGCTGCTTTCT 479
 Db 61 LeuLysSerMetThrAspIleTyrLeuLeuAsnLeuAlaIleSerAspLeuPhePheLeu 80
 QY 480 CTTACTGTCCTTCTGGGCTCACTATGCTGCGCCGCGGAGCTTTGGAATACAATG 539
 Db 81 LeuThrValProPheThrAlaHisTyrAlaAlaAlaGlnTrpAspPheGlyAsnThrMet 100
 QY 540 TGTCAACTCTTGACAGGCTCTATTTATAGGCTTCTCTCTGGAATCTCTCATCATC 599
 Db 101 CysGlnLeuLeuThrGlyLeuTyrPheIleGlyPheSerGlyIlePhePheIleIle 120
 QY 600 CTCCTGACAAATCGATAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 659
 Db 121 LeuLeuThrIleAspArgTyrLeuAlaValAlaHisAlaValPheAlaLeuLysAlaArg 140
 QY 660 ACGTCCACCTTTGGGTTGGTGACAAAGTGTGATGATGATGATGATGATGATGATGATG 719
 Db 141 ThrValThrPheGlyValValThrSerValIleThrTrpValValAlaValPheAlaSer 160
 QY 720 CTCACAGGATCATCTTACAGATCTCAAAAGAGGCTTCTTATACCTGACCTGACCTGCT 779
 Db 161 LeuProGlyIleIleIlePheThrArgSerGlnLysGlnGlyLeuHisTyrThrCysSer 180
 QY 780 CATTTTCCATAC-----ATTAAGATAGTCAT 806
 Db 181 HisPheProTyrSerGlnTyrGlnPheThrPlysAsnPheGlnThrLeuLysIleValI 200
 QY 807 CTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 866
 Db 200 eLeuGlyLeuValLeuProLeuLeuValMetValIleCysTyrSerGlyIleLeuLysTh 220
 QY 867 TCTGCTTCGCTCGAATAG 926
 Db 220 rLeuLeuArgCysArgAsnGlnLysLysArgHisArgAlaValArgLeuIlePheThrI 240
 QY 927 CATGATTGTTTATTTTCTTCTGCGGCTCCCTACACATGTCCTTCTCTCTGACACCTT 986

CC This is the amino acid sequence of wild-type human CCR5, which
 CC serves as a co-receptor for infection by macrophage-tropic
 CC (M-tropic) strains of HIV-1. The invention relates to the
 CC identification of a CCR5 variant (see AAW88231), designated CCR5m303,
 CC comprising the first two transmembrane domains of wild-type CCR5,
 CC but lacking transmembrane domains 3-7. The presence of the
 CC CCR5m303 variant with the wild type CCR5 allele shows a positive
 CC correlation with resistance to infection with M-tropic HIV-1

Db 240 eMetIleValTyrPheLeuPheThrAlaProTyrAsnIleValLeuLeuLeuAsnThrPh 260
 QY 987 CGAGAAATCTTTGGCTTGAATATTCAGTAGCTCTACAGAGTTGGACCAAGTATGCA 1046
 Db 260 eGlnGluPheGlyLeuAsnAsnCysSerSerAsnArgLeuAspGlnAlaMetG1 280
 QY 1047 GTGACAGAGACTCTTGGATGACGACTGCTGCATCAACCCATCATCTATGCTTTGT 1106
 Db 280 nValThrGluThrLeuGlyMetThrHisCysCysIleAsnProIleIleTyrAlaPheVa 300
 QY 1107 CGGGAGAGACTTCAGAACTACCTTCTAGCTTCTTCCTCAAAAGCACATGGCAACGCTT 1166
 Db 300 lGlyGluLysPheArgAsnTyrLeuLeuValPhePheGlnLysHisIleAlaLysArgPh 320
 QY 1167 CTGCAATCTGCTCTATTTCCAGCAAGAGCTCCGAGGAGCAAGCTCAGTTTACAC 1226
 Db 320 eCysLysCysSerIlePheGlnGlnGluAlaProGluArgAlaSerSerValTyrTh 340
 QY 1227 CGATCCACTGGGGAGCAGGAATATCTGTGGGCTTG 1263
 Db 340 rArgSerThrGlyGluGlnGluIleSerValGlyLeu 352

RESULT 6
 AAG80111
 ID AAG80111 standard; Protein; 352 AA.
 AC AAG80111;
 XX 17-JAN-2002 (first entry)
 DT Human CCR5 protein.
 DE
 KW Chemokine; tumour diagnosis; colorectal; prostatic; organ rejection;
 KW inflammation; autoimmune disease; metastasis; bronchial asthma; lupus;
 KW chronic bowel inflammation; rheumatoid arthritis; cytostatic;
 KW antiinflammatory; antischismatic; immunosuppressive; dermatological;
 KW antirheumatic; antiarthritic.
 XX Homo sapiens.
 OS
 XX W0200172830-A2.
 PN
 XX 04-OCT-2001.
 PD
 XX 02-APR-2001; 2001WO-EP03708.
 PF
 XX 31-MAR-2000; 2000DE-1016013.
 PR
 XX (IPFP-) IPF PHARM GMBH.
 PA (FORS/) FORSMANN U.
 XX
 XX Forssmann W, Adermann K, Heitland A, Spodsborg N;
 XX WPI; 2001-626256/72.
 DR
 XX Diagnostic agent containing two or more receptor-specific ligands,
 PT useful for detecting tumors, inflammation etc., also therapeutic use of
 PT ligand inhibitors -
 PS
 XX Disclosure; Page 10; 26pp; German.

CC This invention describes a novel diagnostic agent (A) comprising at least
 CC two different ligands (I) for receptors (II) that are implicated in
 CC disease. (A) are used for the diagnosis of tumors (especially colorectal
 CC or prostatic), organ rejection, inflammation and autoimmune diseases.
 CC Also inhibitors of (I) are used therapeutically against tumors (and their
 CC metastases), inflammation (particularly bronchial asthma or chronic bowel
 CC inflammation), or autoimmune diseases (rheumatoid arthritis or lupus),
 CC where the (cardio)vascular, lymphatic, respiratory, nervous, digestive,
 CC endocrine, motor or urogenital systems or skin are affected, and bone
 CC marrow diseases. The products of the invention are chemokine derivatives
 CC which have cytostatic, antiinflammatory, antiasthmatic,
 CC immunosuppressive, dermatological, antirheumatic, antiarthritic.

CC Chemokines act on specific tumor and inflammatory cells through a
 CC constellation of chemokine receptors (CR), which control migration and
 CC proliferation of these cells. AAG80045-AAG80128 represent human chemokine
 CC fragments used to illustrate the method of the invention.

XX Sequence 352 AA:
 SQ
 Alignment Scores:
 Pred. No.: 338e-187 Length: 352
 Score: 1745.50 Matches: 340
 Percent Similarity: 96.60% Conservative: 1
 Best Local Similarity: 96.32% Mismatches: 0
 Query Match: 66.42% Indels: 12
 DB: 22 Gaps: 1

US-09-938-703-3 (1-1442) x AAG80111 (1-352)

QY 240 ATGATTTCAAGTGTCAAGTCCAAATCATGACATCAATATTATACATCGGAGCCCTGC 299
 Db 1 MetAspTyrGlnValSerSerProIleTyrAspIleAsnTyrThrSerGluProCys 20
 QY 300 CAAAAAATCAATGTGAAGCAAAATCCAGCCCGCTCTCTCGCTCTACTCTACTCTGCTG 359
 Db 21 GlnLysIleAsnValLysGlnIleAlaAlaArgLeuLeuProLeuTyrSerLeuVal 40
 QY 360 TTCATCTTTGTTTGTGGCAACATGCTGCTCATCTCTCATCTGATAAACTGCAAAAGG 419
 Db 41 PheIlePheGlyPheValGlyAsnMetLeuValIleLeuIleAsnCysLysArg 60
 QY 420 CTGAGAGCATGACTGACATCTACCTGTCTCAACCTGCCATCTCTGACCTGTTTTTCCTT 479
 Db 61 LeuLysSerMetThrAspIleTyrLeuLeuAsnLeuAlaIleSerAspLeuPheLeu 80
 QY 480 CTTACTGTCCCTTCTGGGTCACTATGCTGCTGCCCGGAGCTTGGAAATACAAAG 539
 Db 81 LeuThrValProPheTrpAlaHisTyrAlaAlaAlaGlnTrpAspPheGlyAsnThrMet 100
 QY 540 TGTCAACTCTTGACAGGGCTCTATTTATAGGCTTCTCTCTGGAATCTTCTTCATCATC 599
 Db 101 CysGlnLeuLeuThrGlyLeuTyrPheIleGlyPhePheSerGlyIlePhePheIleIle 120
 QY 600 CTCTGACAAATCGATAGTACGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 659
 Db 121 LeuLeuThrIleAspArgTyrLeuAlaValValHisAlaValPheAlaLeuLysAlaArg 140
 QY 660 ACGTCACTTTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 719
 Db 141 ThrValThrPheGlyValValThrSerValIleThrTrpValValAlaValPheAlaSer 160
 QY 720 CTCCAGGAATCATCTTTTACCAGATCTCAAAAGAAGGTCTTCTTATACCTGACCTCT 779
 Db 161 LeuProGlyIleIlePheThrArgSerGlnLysGluGlyLeuHisTyrThrCysSerSer 180
 QY 780 CATTTTCCATAC-----ATTAAGATAGTCAT 806
 Db 181 HisPheProTyrSerGlnTyrGlnPheTrpLysAsnPheGlnThrLeu-LysIleValI1 200
 QY 807 CTTGGGGCTGTGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 866
 Db 200 eLeuGlyLeuValLeuProLeuLeuValMetValIleCysTyrSerGlyIleLeuLysTh 220
 QY 867 TCTGCTCGGTCTCGAATGAGAAGAGAGGACAGGCTGTGAGGCTTATCTTACCATT 926
 Db 220 rLeuLeuArgCysArgAsnGluLysLysArgHisArgAlaValArgLeuIlePheThrI1 240
 QY 927 CATGATGTTTATTTTCTCTCTGGGCTCCCTCAACATGCTCTCTCTCTGACACCTT 986
 Db 240 eMetIleValTyrPheLeuPheThrAlaProTyrAsnIleValLeuLeuLeuAsnThrPh 260
 QY 987 CCAGGAATCTTTGCTGCTGATAATATTGACGTAGCTCTAACAGGTTGGACCAAGCTATGCA 1046
 Db 260 eGlnGluPhePheGlyLeuAsnAsnCysSerSerAsnArgLeuAspGlnAlaMetG1 280

QY 1047 GGTGACAGACACTCTGGGATGAGGACCTGCTGATCAACCCCATCATCTATGCTTGT 1106
 Db |||||||
 280 nValThrGluThrLeuGlyMetThrHisCysCysIleAsnProIleIleTyrAlaPheVa 300
 QY 1107 CGGGAGAGAGTTCAGAACTACCTCTAGTCTTCTTCCAAAGCACATTCGCAACGCTT 1166
 Db |||||||
 300 lGlyGluPheArgAsnTyrLeuLeuValPheGlnLysHisIleAlaLysArgPh 320
 QY 1167 CTCGAAATGCTGTCTTATTTCCAGCAAGAGGCTCCGAGCGAGCAAGCTCAGTTTACAC 1226
 Db |||||||
 320 eCysLysCysCysSerIlePheGlnGlnGluAlaProGluArgAlaSerSerValTyrTh 340
 QY 1227 CGGATCCACTGGGGACGAGGAAATCTCTGTGGCTTG 1263
 Db |||||||
 340 rArgSerThrGlyGluGlnGluIleSerValGlyLeu 352

RESULT 7

AAB82948
 ID AAB82948 standard; Protein; 352 AA.

XX AC AAB82948;

XX DT 21-DEC-2001 (first entry)

XX DE Human HIV-1 co-receptor CCR5.

XX KW CCR5; chemokine; co-receptor; human immunodeficiency virus type 1;

XX KW HIV-1; infection; therapy; vaccine; anti-HIV-1.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Blinding-site 2..18 /note= "binds to HIV-1 gp120"

XX FT WO200164710-A2.

XX PN 07-SEP-2001.

XX PF 28-FEB-2001; 2001WO-US06699.

XX PR 29-FEB-2000; 2000US-185667P.

XX PR 19-MAY-2000; 2000US-205839P.

XX PR 07-FEB-2001; 2001US-267231P.

XX XX (PROG-) PROGENICS PHARM INC.

PA (AARO-) AARON DIAMOND AIDS RES CENT.

XX PI Dragic T, Olson WC;

XX DR WPI; 2001-611273/70.

XX DR N-PSDB; AAB26903.

XX PT Novel compounds comprising specific amino acids within CCR5 (HIV 1
 PT co-receptor) amino terminal domain including negatively charged and two
 PT sulfated tyrosine residues is useful for treating HIV infection in
 PT humans

XX PS Claim 1; Page 30; 163pp; English.

XX CC The present sequence is that of human HIV-1 co-receptor CCR5.

XX CC Amino acids 2-18 in the N-terminal region of CCR5 comprise an HIV-1

XX CC gp120-binding site that determines the specificity of the

XX CC interaction between CCR5 and HIV-1 gp120. Post-translational

XX CC sulfation of the tyrosine residues in the CCR5 N-terminus is

XX CC required for gp120 binding and may critically modulate the

XX CC susceptibility of target cells to HIV-1 infection in vivo. The

XX CC invention provides claimed sulfated peptides (see AAB82947) that

XX CC are based on the CCR5 N-terminal region and which are effective

XX CC for inhibiting HIV-1 binding to CCR5. These peptides are used in

XX CC claimed methods of inhibiting HIV infection of CD4+ cells, of

XX CC preventing CD4+ cells from becoming infected with HIV, of treating

XX CC a subject whose CD4+ cells are infected with HIV, and of

CC identifying an agent which inhibits binding of a CCR5 ligand to a
 CC CCR5 receptor. The methods may be carried out in a subject,
 CC especially a human, infected (therapeutic method), not infected
 CC with HIV (prophylactic method), or in a subject who is not infected
 CC with, but has been exposed to, HIV.

XX SQ Sequence 352 AA;

Alignment Scores:

Pred. No.: 338e-187 Length: 352
 Score: 1745.50 Matches: 340
 Percent Similarity: 96.60% Conservative: 1
 Best Local Similarity: 96.32% Mismatches: 0
 Query Match: 66.42% Indels: 12
 DB: 22 Gaps: 1

US-09-938-703-3 (1-1442) x AAB82948 (1-352)

QY 240 ATGGATATCAAGTGTCAAGTCCCAATCTATGACATCAATATTATATACATCGAGCCCTGC 299

Db 1 MetAspTyrGlnValSerSerProIleTyrAspIleAsnTyrThrSerGluProCys 20

QY 300 CAAAAATCAATGTCAAGCAAAATCGACCGCGCTCTCCCTCCGCTCTACTCATCTGTGTG 359

Db 21 GlnLysIleAsnValLysGlnIleAlaAlaArgLeuLeuProLeuTyrSerLeuVal 40

QY 360 TTCACTTTGGTTTGGGCAACATGCTGTATCTCTCATCTCTGATATAAATCAAGG 419

Db 41 PheIlePheGlyPheValGlyAsnMetLeuValIleLeuIleAsnCysLysArg 60

QY 420 CTGAAGACGATGACATCTACCTGTCTCAACCTGCCATCTCTGACCTGTGTTTCCTT 479

Db 61 LeuLysSerMetThrAspIleTyrLeuLeuAsnLeuAlaIleSerAspLeuPhePheLeu 80

QY 480 CTTACTGTCCTCTCTGGGCTCACTATGCTCCGCCAGTGGGACTTTGGAATACAAAG 539

Db 81 LeuThrValProPheThrAlaHisTyrAlaAlaAlaGlnTrpAspPheGlyAsnThrMet 100

QY 540 TGTCAACTCTTGACAGGCTCTATTTATAGGCTTCTCTCTGGAATCTTCTTCATCATC 599

Db 101 CysGlnLeuLeuThrGlyLeuTyrPheIleGlyPheSerGlyIlePhePheIleIle 120

QY 600 CTCTGACAAATCGATAGGTACCTGGCTGTCTCCATGCTGTGTTGCTTTAAAGCCAG 659

Db 121 LeuLeuThrIleAspArgTyrLeuAlaValHisAlaValPheAlaLeuLysAlaArg 140

QY 660 ACGGTACCTTTGGGTGGTGACAGTGTGATCATCTGGTGGTGGCTGTGTTGGCTCT 719

Db 141 ThrValThrPheGlyValValThrSerValIleThrTrpValValAlaValPheAlaSer 160

QY 720 CTCGAGGAATCATCTTTACGATCTCAAAAAGAGGTCTTCATTACACCTGCAGCTCT 779

Db 161 LeuProGlyIleIlePheThrArgSerGlnLysGluGlyLeuHisTyrThrCysSerSer 180

QY 780 CATTTCCTCATAC-----ATTAAGATAGTGTAT 806

Db 181 HisPheProTyrSerGlnTyrGlnPheTrpLysAsnPheGlnThrLeu-LysIleValIle 200

QY 807 CTTGGGSGCTGCTCGCGCTCTGTCTGATGCTCATCTGCTACTCTCGGAATCTTAAAAAC 866

Db 200 eLeuGlyLeuValLeuProLeuLeuValMetValIleCysTyrSerGlyIleLeuLysTh 220

QY 867 TCTGCTCGGTCTCGAAATGAGAAGAGGACAGGCTGTGAGGCTTATCTTCACCAT 926

Db 220 rLeuLeuArgCysArgAsnGlnLysLysArgHisArgAlaValArgLeuIlePheThrIle 240

QY 927 CATGATTGTTTATTTCTTCTTGGGCTCCCTACAAATGTCCTTCTCTGACACCTT 986

Db 240 eMetIleValTyrPheLeuPheThrAlaProTyrAsnIleValLeuLeuLeuAsnThrPh 260

QY 987 CCAGGAATCTTTGGCTCGAATATTTGCAAGTAGCTCTTAACAGCTTGGACCAAGCTATGCA 1046

Db 260 eGlnGluPhePheGlyLeuAsnAsnCysSerSerSerAsnArgLeuAspGlnAlaMetGln 280

Db 300 lGlyluLysPheArgAsnTyrLeuLeuValPhePheGlnLysHisIleAlaLysArgPh 320
 QY 1167 CTGCAATGCTGTTCTATTTCCAGCAAGAGGCTCCGAGCGAGCAAGCTCAGTTACAC 1226
 Db 320 eCysLysCysSerIlePheGlnGlnGlnAlaProGluArgAlaSerValTyrTh 340

QY 1227 CCGATCCACTGGGAGCAGGAATATCTGTGGGCTG 1263

Db 340 rArgSerThrGlyGlnGlnGlnIleSerValGlyLeu 352

RESULT 9

AAE04321
 ID AAE04321 standard; Protein; 352 AA.

XX AC AAE04321;

XX AC AAE04321;

XX 04-SEP-2001 (first entry)

XX Human chemokine receptor (CCR), CC-CR-5 related protein #2.

XX Human; transformed mammalian cell; CD4; reporter gene; translocation;
 KW human immuno deficiency virus; HIV; long terminal repeat; LTR; therapy;
 KW chemokine receptor; CCR; cellular dysfunction; HIV infection; cofactor;
 KW CC-CR-5; envelope glycoprotein; anti-HIV.

XX Homo sapiens.

XX US6258527-B1.

XX 10-JUL-2001.

XX 21-MAY-1997; 97US-0861105.

XX 20-MAY-1996; 96US-0017157.

XX 19-JUN-1996; 96US-0020043.

XX 19-MAY-1997; 97US-0858660.

XX (AARO-) ARON DIAMOND AIDS RES CENT.

XX (DTNY) UNIV NEW YORK STATE.

PI Littman DR, Deng H, Ellmeier W, Landau NR, Liu R;

XX WPI; 2001-417127/44.
 XX N-PSDB; AAD08577.

XX Transformed mammalian cell (1) that contains a CD4 gene, reporter gene
 PT and HIV LTR for identification of drugs and antibodies for treatment of
 PT HIV -

XX Disclosure; Column 47-50; 37pp; English.

XX The present invention relates to a transformed mammalian cell that
 CC contains a gene encoding CD4, a construct encoding a reporter gene under
 CC the regulation of an human immuno deficiency virus (HIV) long terminal
 CC repeat (LTR) and that has been transduced with a vector encoding a
 CC human chemokine receptor (CCR) where the CD4 and the CCR are present on
 CC the cell surface of transformed mammalian cell. The invention is useful
 CC for identifying drugs or antibodies that interfere with the
 CC translocation of HIV into transformed mammalian cell or for identifying
 CC a human chemokine receptor that facilitates the infection of a
 CC particular HIV strain into the transformed mammalian cell. Compounds
 CC identified can be used to treat cellular dysfunction and to prevent or
 CC combat HIV infection. The present sequence is a human chemokine receptor
 CC (CCR), CC-CR-5 related protein. CC-CR-5 is the principal cofactor for
 CC entry mediated by the envelope glycoproteins of primary macrophage-tropic
 CC strains of HIV-1.

XX Sequence 352 AA;

SQ

Alignment Scores: 3.38e-187 Length: 352

Pred. No.: 1745.50 Matches: 340

Score:

Percent Similarity: 96.60% Conservative: 1
 Best Local Similarity: 96.32% Mismatches: 0
 Query Match: 66.42% Indels: 12
 DB: 22 Gaps: 1

US-09-938-703-3 (1-1442) x AAE04321 (1-352)

QY 240 ATGGATTATCAAGTGTCAAGTCCAAATCTATGACATCAATATTATATACATCGAGCCCTGC 299
 Db 1 MetAspTyrGlnValSerProIleTyrAspIleAsnTyrTyrThrSerGluProCys 20
 QY 300 CAARAAATCAATGTGAAGCAATCGAGCCGCTCCCTCGCTCGCTACTACTACTACTGCTG 359
 Db 21 GlnLysIleAsnValLysGlnIleAlaAlaArgLeuLeuProIleTyrSerLeuVal 40
 QY 360 TTCATCTTTTGGTTTGTGGCAACATCTGCTCATCTCTCTCATCTCTGTAACATCAAAAGG 419
 Db 41 PheIlePheGlyPheValGlyAsnMetLeuValIleLeuIleLeuIleAsnCysLysArg 60
 QY 420 CTGAAGACATGACTGACATCTACTGCTCAACCTGGCCATCTCTGACCTGTTTTCCTT 479
 Db 61 LeuLysSerMetThrAspIleTyrLeuLeuAsnLeuAlaIleSerAspLeuPheLeu 80
 QY 480 CTTACTGTCCCTCTCTGGGCTCACTATCTCGCCGCCAGTGGGACTTTGGAATACAATG 539
 Db 81 LeuThrValProPheTyrAlaHisTyrAlaAlaGlnTyrAspPheGlyAsnThrMet 100
 QY 540 TGTCAACTCTTGACAGGCTCTATTTATAGGCTTCTCTCTGGAATCTCTTCATCATC 599
 Db 101 CysGlnLeuLeuThrGlyLeuTyrPheIleGlyPheSerGlyIlePhePheIleIle 120
 QY 600 CTCTCGAATCAATGATAGTACCTGCTGCTCATCTGCTGTTGTTTAAAGCCAGG 659
 Db 121 LeuLeuThrIleAspArgTyrLeuAlaValAlaHisAlaValPheAlaLeuLysAlaArg 140
 QY 660 ACGGTCACTTTGGGTGGTGAGTGTGATCCTGCTGGTGGTGGTGGTGGTGGTGGTGGT 719
 Db 141 ThrValThrPheGlyValValThrSerValIleTyrTyrValValAlaValPheAlaSer 160
 QY 720 CTCACGAGCAATCACTTACCAGATCTCAAAAGAGGTCTTCATTCACCTGACCTCT 779
 Db 161 LeuProGlyIleIlePheThrArgSerGlnLysGluGlyLeuHisTyrThrCysSer 180
 QY 780 CATTTCCTATAC-----ATTAAGATAGTCAAT 806
 Db 181 HisPheProTyrSerGlnTyrGlnPheTyrLysAsnPheGlnThrLeuLysIleVal 200
 QY 807 CTTGGGGCTGGTCTGCGGCTGCTGTGATGGTCACTGCTACTCGGGAATCTCAAAAC 866
 Db 200 eLeuGlyLeuValLeuProLeuValMetValIleCysTyrSerGlyIleLeuLysTh 220
 QY 867 TCTGCTCGGTGTCGAATCAGAACGACGACGAGGCTGTGAGGCTTATCTTCACCAT 926
 Db 220 rLeuLeuArgCysArgAsnGlnLysLysArgHisArgAlaValArgLeuIlePheThr 240
 QY 927 CATGATTTGTTATTTTCTCTCTGGGCTCCCTACAACATGTCCTCTCTGACACCTT 986
 Db 240 eMetIleValTyrPheLeuPheTyrAlaProTyrAsnIleValLeuLeuLeuAsnThrPh 260
 QY 987 CCAGGAATTTTGGCTGGAATTAATGAGTGTCTTACAGTTGGACCAAGCTATGCA 1046
 Db 260 eGlnGluPhePheGlyLeuAsnAsnCysSerSerSerAsnArgLeuAspGlnAlaMetG 280
 QY 1047 GGTGACAGAGACTCTTGGGATGACGACTGCTGATCAACCCATCATCTATCCCTTGT 1106
 Db 280 nValThrGluThrLeuGlyMetThrHisCysCysIleAsnProIleIleTyrAlaPhe 300
 QY 1107 CGGGAGAGAGTTCAGAACTACCTCTTAGTCTTCTTCCAAAAGACATTCGCAACGCTT 1166
 Db 300 lGlyGluLysPheArgAsnTyrLeuValPhePheGlnLysHisIleAlaLysArgPh 320
 QY 1167 CTGCAATGCTGTTCTATTTTCAGCAAGGCTCCGAGCGAGCAAGCTCACTTTACAC 1226

Db 320 eCysLysCysSerIlePheGlnGlnGlnAlaProGluArgAlaSerValTyrTh 340

QY 1227 CCGATCCACTGGGAGCAGAAATATCTGTGGCTTG 1263

Db 340 rArgSerThrGluGlnGlnIleSerValGlyLeu 352

RESULT 10

AB08343

ID AB08343 standard; Protein; 352 AA.

XX AC AB08343;

XX 18-JUN-2002 (first entry)

XX Human chemokine (C-C motif) receptor 5 polypeptide.

XX Human chemokine C-C motif receptor 5; CCR5; haplotype pair; isogene;
 KW single nucleotide polymorphism; SNP; human immunodeficiency virus 1;
 KW HIV-1 infection; acquired immunodeficiency syndrome; AIDS; antiviral;
 KW genotype; polymorphic variant; transgenic; drug screening; gene therapy;
 KW chromosome 3p21.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Misc-difference 55

FT /label= Leu, Gln

FT Misc-difference 182

FT /label= Phe, Leu

FT Misc-difference 223

FT /label= Arg, Gln

XX WC200177125-A2.

XX 18-OCT-2001.

XX 04-APR-2001; 2001WO-US10708.

XX 05-APR-2000; 2000US-194361P.

XX (GENA-) GENA1SSANCE PHARM INC.

XX Choi JY, Kitem SE, Koshiy B;

XX WPI; 2002-041282/05.

XX N-PSDB; ABA97318, ABA97319.

XX New haplotypes of the human chemokine (C-C motif) receptor 5 gene,
 useful to diagnose and treat diseases associated with its abnormal
 expression or function, including human immunodeficiency virus-1
 infection -

PS Claim 29; Fig 3; 61pp; English.

XX The present sequence is that of a polypeptide encoded by the human
 CC chemokine (C-C motif) receptor 5 (CCR5) gene of the invention [see
 CC ABA97319, ABA97319]. The specification describes haplotyping the CCR5
 CC gene of an individual by determining if the individual has one of the
 CC CCR5 haplotypes or haplotype pairs fully defined in the specification.
 CC The specification also describes an isolated polynucleotide comprising a
 CC nucleotide sequence which is a polymorphic variant of the reference CCR5
 CC gene sequence and comprises an isogene defined by a haplotype described
 CC in the specification and its encoded polypeptide. The methods of the
 CC invention are useful to diagnose and develop treatment for diseases
 CC associated with abnormal expression or function of the gene. The CCR5
 CC isogenes and the screened compounds are useful for treating human
 CC immunodeficiency virus (HIV)-1 infection and the progression to acquired
 CC immunodeficiency syndrome (AIDS). The invention has antiviral
 CC applications. The specification describes genotyping the CCR5 gene of an
 CC individual; predicting a haplotype pair for the CCR5 gene of an
 CC individual; identifying an association between a trait and a haplotype or
 CC haplotype pair of the CCR5 gene. The specification describes a
 CC composition comprising a genotyping oligonucleotide for detecting a CCR5

CC polymorphism: a recombinant non-human organism transformed with CCR5
 CC polynucleotide expressing a CCR5 protein encoded by the variant sequence;
 CC an isolated antibody specific for the CCR5 polypeptide and a method for
 CC screening drugs targeting the CCR5 polypeptide.

XX Sequence 352 AA;

Alignment Scores:

Pred. No.: 3.38e-187 Length: 352
 Score: 1745.50 Matches: 340
 Percent Similarity: 96.60% Conservative: 1
 Best Local Similarity: 96.32% Mismatches: 0
 Query Match: 66.42% Indels: 12
 DB: Gaps: 23

US-09-938-703-3 (1-1442) x ABB08343 (1-352)

QY 240 ATGGATTATCAAGTGTCAAGTCCAAATCTATGACATCAATATTATATACATCGAGCCCTGC 299

Db 1 MetAspTyrGlnValSerSerProIleTyrAspIleAsnTyrTyrThrSerGluProCys 20

QY 300 CAAAAATCAATGTGAAGCAATCGAGCCGCCCTCCCTCGCTCGCTCTACTCTACTGGTG 359

Db 21 GlnTyrIleAsnValLysGlnIleAlaAlaArgLeuLeuProLeuTyrSerLeuVal 40

QY 360 TTCATCTTGTGTTTGTGGGCAACATCTGCTCATCTCTCATCTCTGATAAATGCAAAAGG 419

Db 41 PheIlePheGlyPheValGlyAsnMetLeuValIleLeuIleLeuIleAsnCysLysArg 60

QY 420 CTGAAGAGCATGACTGACATCTACTCTCAACCTGGCCATCTCTGACCTGTTTCTCTT 479

Db 61 LeuLysSerMetThrAspIleTyrLeuLeuAsnLeuAlaIleSerAspLeuPheLeu 80

QY 480 CTTACTGTCCCTCTGGGCTCAGTATGCTGCCGCCAGTGGGACTTGGGAATCAATG 539

Db 81 LeuThrValProPheTrpAlaHisTyrAlaAlaIleGlnTrpAspPheGlyAsnThrMet 100

QY 540 TGTCAACTCTTGACAGGCTCTATTATAGGCTTCTCTCTGGAATCTTCTTCATCATC 599

Db 101 CysGlnLeuLeuThrGlyLeuTyrPheIleGlyPheSerGlyIlePhePheIleIle 120

QY 600 CTCCTGACAATCGATAGGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 659

Db 121 LeuLeuThrIleAspArgTyrLeuAlaValValHisAlaValPheAlaLeuLysAlaArg 140

QY 660 ACGTCACTTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 719

Db 141 ThrValThrPheGlyValValThrSerValIleThrTrpValValAlaValPheAlaSer 160

QY 720 CTCACGAGGAATCATCTTTACAGATCTCAAAAGAGGCTTTCATTAACCTGCACTCT 779

Db 161 LeuProGlyIleIlePheThrArgSerGlnLysGluGlyLeuHisTyrThrCysSer 180

QY 780 CATTTTCCATAC-----ATTAAGATAGTCAAT 806

Db 181 HisPheProTyrSerGlnTyrGlnPheTrpLysAsnPheGlnThrLeu-LysIleValI 200

QY 807 CTTGGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 866

Db 200 eLeuGlyLeuValLeuProLeuLeuValMetValIleCysTyrSerGlyIleLeuLysTh 220

QY 867 TCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 926

Db 220 rLeuLeuArgCysArgAsnGlnLysLysArgHisArgAlaValAlaArgLeuIlePheThrI 240

QY 927 CATGATGTTTATTTCTCTCTCTGGGCTCCCTACACATGTCTCTCTCTGACACCTT 986

Db 240 eMetIleValTyrPheLeuPheTrpAlaProTyrAsnIleValLeuLeuLeuAsnThrPh 260

QY 987 CCAGGATCTTTGGCTGATTAATTCAGTAGCTCTAACAGGTGGACCAAGCTATGCA 1046

Db 260 eGlnGluPhePheGlyLeuAsnAsnCysSerSerAsnArgLeuAspGlnAlaMetG 280

QY 1047 GGTGACAGAGACTCTTGGATGACGACTGCTGCATCAACCCCATCATCTATGCGCTTGT 1106
 XX
 Db 280 nValThrGluThrLeuGlyMetThrHisCysCysLeuAsnProIleIleTyrAlaPheVa 300

QY 1107 CGGGGAGAGTTCAGAACTACCTCTAGTCTTCTCCAAAGCAGCATGCCAAAGCTTT 1166
 XX
 Db 300 LGlyGluLysPheArgAsnTyrLeuLeuValPhePheGlnLysHisAlaLysArgPh 320

QY 1167 CTGCAATGCTGTTCTATTTTCCAGCAAGAGCTCCGAGCGAGCAAGCTCAGTTTACAC 1226
 XX
 Db 320 eCysLysCysSerIlePheGlnGlnGlnAlaProGluArgAlaSerSerValTyrTh 340

QY 1227 CCGATCACTGGGGAGCAGGAATATCTCTGGGGTTG 1263
 XX
 Db 340 rArgSerThrGlyGluGlnGluIleSerValGlyLeu 352

RESULT 11
 AAM52828
 ID AAM52828 standard; Protein; 352 AA.
 XX
 AC AAM52828;
 XX
 DT 22-FEB-2002 (first entry)
 XX
 DE Human CC chemokine receptor 5 (CCR5).
 XX
 KW CCR5; CC chemokine receptor 5; human; HIV infection;
 KW human immunodeficiency virus; AIDS; acquired immunodeficiency syndrome;
 KW drug screening; identification.
 XX
 OS Homo sapiens.
 XX
 PN W0200171346-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 21-MAR-2001; 2001WO-US09155.
 XX
 PR 21-MAR-2000; 2000US-190946P.
 PR 21-MAR-2000; 2000US-190996P.
 PR 21-MAR-2000; 2000US-191299P.
 PR 20-MAR-2001; 2001US-0813448.
 PR 20-MAR-2001; 2001US-0813651.
 PR 20-MAR-2001; 2001US-0813653.
 XX
 PA (CONS-) CONSENSUS PHARM INC.
 XX
 PI Nestor JJ, Wilson CJ, See RH, Tan Hehir CA;
 XX
 DR WPI: 2002-010610/01.
 DR N-PSDB; ARA02317.
 XX
 XX Identifying CC chemokine receptor 5 binding compound for treating AIDS,
 PT comprises binding a molecule from library to a molecule having binding
 PT property corresponding to CCR5 and identifying bound molecule -
 XX
 XX Example 3; Fig 4a; 50pp; English.
 XX
 XX The invention relates to a method for identifying a binding compound
 CC for CC chemokine receptor 5 (CCR5). The method involves screening a
 CC library of test molecules (particularly peptides) with immobilised CCR5,
 CC and then identifying those molecules which bind. The invention also
 CC relates to CCR5-binding molecules identified using the method of the
 CC invention, methods for identifying consensus motifs for CCR5-binding
 CC peptides, a transfer vector encoding tagged CCR5, a computer-aided
 CC methods for determining the relative binding affinity of a test molecule
 CC to CCR5 and a computer aided drug screening assay that utilises the
 CC three-dimensional structure of CCR5. Compounds identified using the
 CC methods of the invention are useful for treating or preventing HIV
 CC (human immunodeficiency virus) infection or AIDS (acquired
 CC immunodeficiency syndrome) in a patient. The methods of the invention
 CC may also be used to identify agonists or antagonists of the interaction
 CC of CCR5 with its natural ligand, and to determine a binding motif for

CC CCR5. The present sequence represents human CCR5.

XX
 SQ Sequence 352 AA;

Alignment Scores:
 Pred. No.: 3,38e-187 Length: 352
 Score: 1745.50 Matches: 340
 Percent Similarity: 96.60% Conservative: 1
 Best Local Similarity: 96.32% Mismatches: 0
 Query Match: 66.42% Indels: 12
 Gaps: 23

US-09-938-703-3 (1-1442) x AAM52828 (1-352)

QY 240 ATGGATTATCAAGTGTCAAGTCCCAATCTATGATCAATATTTATACATCGAGCCCTGC 299
 XX
 Db 1 MetAspTyrGlnValSerSerProIleTyrAspIleAsnTyrTyrSerGluProCys 20

QY 300 CAAAAATCAATGTGAAGCAAAATGCAGCCGCTCTGCTCGCTCTACTACTACTGCTGTG 359
 XX
 Db 21 GlnLysIleAsnValLysGlnIleAlaAlaArgLeuLeuProIleTyrSerLeuVal 40

QY 360 TTCATCTTTTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 419
 XX
 Db 41 PheIlePheGlyPheValGlyAsnMetLeuValIleLeuIleLeuLeuAsnCysLysArg 60

QY 420 CTGAGAGACATGACTGACATCTACCTGCTCACTGGCCATCTCTGACCTGTGTTTCTT 479
 XX
 Db 61 LeuLysSerMetThrAspIleTyrLeuLeuAsnLeuAlaIleSerAspLeuPheLeu 80

QY 480 CTTACTGTCCCTCTCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 539
 XX
 Db 81 LeuThrValProPheThrPalaHisTyrAlaAlaGlnTyrPaspPheGlyAsnThrMet 100

QY 540 TGTCACACTCTTGACAGGGCTCTATTTTATAGGCTTCTCTCGGAATCTTCTTCATCATC 599
 XX
 Db 101 CysGlnLeuLeuThrGlyLeuTyrPheIleGlyPhePheSerGlyIlePhePheIleIle 120

QY 600 CTCCTGACATCGATAGTACCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 659
 XX
 Db 121 LeuLeuThrIleAspArgTyrLeuAlaValHisAlaValPheAlaLeuLysAlaArg 140

QY 660 ACGGTCACTCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 719
 XX
 Db 141 ThrValThrPheGlyValValThrSerValIleThrTyrValAlaValPheAlaSer 160

QY 720 CTCACAGGAATCATCTTACAGACTCTCAAAAGAGGCTCTTCATACACCTGACCTCTCT 779
 XX
 Db 161 LeuProGlyIleIlePheThrArgSerGlnLysGlnGlyLeuHisTyrThrCysSerSer 180

QY 780 CATTTCCTCATAC-----ATTAAGAGATAGTCAT 806
 XX
 Db 181 HisPheProTyrSerGlnTyrGlnPheThrLysAsnPheLeuThrLeuLysIleValI 200

QY 807 CTTGGGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 866
 XX
 Db 200 eLeuGlyLeuValLeuProLeuLeuValMetValIleCysTyrSerGlyIleLeuLysTh 220

QY 867 TCTGCTTCGCTGCTCAATGAG 926
 XX
 Db 220 rLeuLeuArgCysArgAsnGlnLysLysArgHisArgAlaValArgLeuIlePheThrI 240

QY 927 CATGATTGTTTATTTTCTCTCTGCGCTCCCTACACATTCCTCTCTCTCTCTCTCTCTCT 986
 XX
 Db 240 eMetIleValTyrPheLeuPheThrPalaProTyrAsnIleValLeuLeuLeuAsnThrPh 260

QY 987 CCAGGAATCTTTGGCTCGCTGAATTAATTCAGTAGTCTTAACAGGTTGGACAGCTATGCA 1046
 XX
 Db 260 eGlnGluPhePheGlyLeuAsnAsnCysSerSerSerAsnArgLeuAspGlnAlaMetG 280

QY 1047 GGTGACAGAGACTCTTGGATGACGACTGCTGCATCAACCCCATCATCTATGCGCTTGT 1106
 XX
 Db 280 nValThrGluThrLeuGlyMetThrHisCysCysIleAsnProIleIleTyrAlaPheVa 300

QY 1107 CGGGGAGAGTTCAGAACTACCTCTTACTCTTCTTCCTCAAAAGACATTCGCAACGCTT 1166
 |||||
 Db 378 IGlyGluLysPheArgAsnTyrLeuLeuValPhePheGlnLysHisIleAlaLysArgPh 398
 |||||
 QY 1167 CTGCAATGCTGTCTATTTCAGCAAGAGGCTCCGAGGAGCAAGCTCAGTTTACAC 1226
 |||||
 Db 398 eCysLysCysSerIlePheGlnGlnGluAlaProGluArgAlaSerValTyrTh 418
 |||||
 QY 1227 CCGATCCACTGGGAGCAGGAATATCTGTGGGCTTG 1263
 |||||
 Db 418 rArgSerThrGlyGluGlnGluIleSerValGlyLeu 430
 |||||

RESULT 13

ABB56342
 ID ABB56342 standard; Protein; 352 AA.

AC ABB56342;

DT 18-FEB-2002 (first entry)

DE Non-endogenous human GPCR protein, SEQ ID NO: 477.

KW Human; G protein-coupled receptor; GPCR; non-endogenous; mutant;
 constitutively activated GPCR; agonist; disease.

OS Homo sapiens.

OS Synthetic.

PN WO2001/7172-A2.

XX 18-OCT-2001.

XX 05-APR-2001; 2001WO-US11098.

XX 07-APR-2000; 2000US-195747P.

PA (AREN-) ARENA PHARM INC.

PI Lehmann-Bruinsma K, Liaw CW, Lin I;

DR WPI; 2001-648759/74.

DR N-PSDB; AB197978.

PT Identifying agonists of G protein-coupled receptors (GPCRs) for use in
 disease treatment, comprises contacting candidate compounds with
 versions of GPCRs -

PS Claim 1; Page 277-278; 394pp; English.

XX The invention relates to G protein-coupled receptors (GPCRs) for which
 the endogenous ligand has been identified. Non-endogenous
 CC constitutively activated versions of known GPCRs are used in the
 CC invention for the direct identification of candidate compounds as
 CC receptor agonists, inverse agonists or partial agonists. Such
 CC agonists are useful as therapeutic agents for diseases or disorders
 CC associated with GPCRs. The present sequence is a non-endogenous
 CC version of a known human GPCR.

SQ Sequence 352 AA;

Alignment Scores:

| Pred. No.: | 1-6e-186 | Length: | 352 |
|------------------------|----------|---------------|-----|
| Score: | 1739.50 | Matches: | 339 |
| Percent Similarity: | 96.32% | Conservative: | 1 |
| Best Local Similarity: | 96.03% | Mismatches: | 12 |
| Query Match: | 66.19% | Indels: | 1 |
| DB: | 22 | Gaps: | 1 |

US-09-938-703-3 (1-1442) x ABB56342 (1-352)

QY 240 ATGATATCAAGTGTCAAGTCCATCTATGACATCAATATTATACATCGGAGCCCTGC 299

Db 1 MetAspTyrGlnValSerSerProIleTyrAspIleAsnTyrThrSerGluProCys 20
 |||||

QY 300 CAAAAATCAATGTGAAGCAATCCAGCCGCTCTCGCTCCGCTCTACTCACTGCTG 359
 |||||
 Db 21 GlnLysIleAsnValLysGlnIleAlaAlaArgLeuLeuProLeuTyrSerLeuVal 40
 |||||
 QY 360 TTCATCTTTGTTTGTGGCAACATGTGTCTCATCTCATCTGATAAACTGCAAAAGG 419
 |||||
 Db 41 PheIlePheGlyPheValGlyAsnMetLeuValIleLeuIleLeuLeuAsnCysLysArg 60
 |||||
 QY 420 CTGAAGAGCATGACTGACATCTACCTGCTCAACCTGGCCATCTCTGACCTGTTCCTT 479
 |||||
 Db 61 LeuLysSerMetThrAspIleTyrLeuLeuAsnLeuAlaIleSerAspLeuPheLeu 80
 |||||
 QY 480 CTTCATCTCCCTTCGCGGCTCACTATCTGCGCCGCTGGGACTTTGCAATACAATG 539
 |||||
 Db 81 LeuThrValProPheThrPalaHisTyrAlaAlaAlaGlnTrpAspPheGlyAsnThrMet 100
 |||||
 QY 540 TGTCAACTCTTGACAGGGCTCTATTATTAGGCTTCTCTCTGGAATCTTCCTTCATCATC 599
 |||||
 Db 101 CysGlnLeuLeuThrGlyLeuTyrPheIleGlyPheSerGlyIlePhePheIleIle 120
 |||||
 QY 600 CTCTGACATCGATAGGTACCTGGCTGTCTGCTCATCTGCTGTGCTTTAAAGCCAGG 659
 |||||
 Db 121 LeuLeuThrIleAspArgTyrLeuAlaValHisAlaValPheAlaLeuLysAlaArg 140
 |||||
 QY 660 ACGSTCACCTTTGGGTGGTGCAAGTGTGATCACTTGGTGTGGCTGTGTTCGCTCT 719
 |||||
 Db 141 ThrValThrPheGlyValValThrSerValIleThrIrpValValAlaValPheAlaSer 160
 |||||
 QY 720 CTCGAGGAATCATCTTTACCATCTCAAAAGAGGTCTTCATTACACCTGACGCTCT 779
 |||||
 Db 161 LeuProGlyIleIlePheThrArgSerGlnLysGluGlyLeuHisTyrThrCysSerSer 180
 |||||
 QY 780 CATTTTCATAC-----ATTAAAGATAGTCAT 806
 |||||
 Db 181 HisPheProTyrSerGlnTyrGlnPheTrpLysAsnPheGlnThrLeu-LysIleValI 200
 |||||
 QY 807 CTGGGGCTGGTCCCTGCTGCTGTCTCATGTCTCATCTCGGAATCTCTAAAC 866
 |||||
 Db 200 eLeuGlyLeuValLeuProLeuLeuValMetValIleCysTyrSerGlyIleLeuLysTh 220
 |||||
 QY 867 TCTGCTTCGGTCTCGAAATGAGAGGACAGGCTGTGAGGCTTATCTTCACCAT 926
 |||||
 Db 220 rLeuLeuArgCysArgAsnGlnLysLysArgHisArgAlaLysArgLeuIlePheThrI 240
 |||||
 QY 927 CATCATTTGTTTATTTCTCTCTGGGCTCCCTACACATTTGCTCTCTCTGGAACCTT 986
 |||||
 Db 240 eMetIleValTyrPheLeuPheThrPalaProTyrAsnIleValLeuLeuLeuAsnThrPh 260
 |||||
 QY 987 CCAGGAATCTTTGGCTGCAATATTCAGTAGCTCTAACAGGTGGACCAAGCTATGCA 1046
 |||||
 Db 260 eGlnGluPhePheGlyLeuAsnAsnCysSerSerSerAsnArgLeuAspGlnAlaMetG 280
 |||||
 QY 1047 GGTGACAGACATCTTGGGATGAGGACGCTGCTGATCAACCCCATCATCTATGCTTGT 1106
 |||||
 Db 280 nValThrGluThrLeuGlyMetThrHisCysCysIleAsnProIleIleTyrAlaPheVa 300
 |||||
 QY 1107 CGGGGAGAAGTTCAGAACTACCTCTTACTCTTCTCAAAAGACATTCGCAACGCTT 1166
 |||||
 Db 300 IGlyGluLysPheArgAsnTyrLeuLeuValPhePheGlnLysHisIleAlaLysArgPh 320
 |||||
 QY 1167 CTGCAATGCTGTCTATTTCAGCAAGAGGCTCCCGAGGAGCAAGCTCAGTTTACAC 1226
 |||||
 Db 320 eCysLysCysCysSerIlePheGlnGlnGluAlaProGluArgAlaSerValTyrTh 340
 |||||
 QY 1227 CCGATCCACTGGGAGCAGGAATATCTGTGGGCTTG 1263
 |||||
 Db 340 rArgSerThrGlyGluGlnGluIleSerValGlyLeu 352
 |||||

RESULT 14

AAW52829

ID AAW52829 standard; Protein; 352 AA.

XX

AC AAM52829;
XX 22-FEB-2002 (first entry)
DT Human CCR5 Gln 55 variant.
XX CCR5; CC chemokine receptor 5; human; HIV infection;
KW human immunodeficiency virus; AIDS; acquired immunodeficiency syndrome;
KW drug screening; identification; variant.
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH Misc-difference 55 /note= "Glu replaces wild-type Leu; encoded by CNG"
FT Misc-difference 58 /note= "Encoded by AGC"
FT
FT
XX WO200171346-A2.
XX
XX 27-SEP-2001.
XX
XX 21-MAR-2001; 2001WO-US09155.
XX
XX 21-MAR-2000; 2000US-190946P.
XX 21-MAR-2000; 2000US-190996P.
XX 21-MAR-2000; 2000US-191299P.
XX 20-MAR-2001; 2001US-081344B.
XX 20-MAR-2001; 2001US-0813651.
XX 20-MAR-2001; 2001US-0813653.
XX
XX (CONS-) CONSENSUS PHARM INC.
XX
XX Nestor JJ, Wilson CJ, See RH, Tan Behir CA;
XX
XX WPI: 2002-010610/01.
XX N-PSDB; ABA02318.
XX
XX Identifying CC chemokine receptor 5 binding compound for treating AIDS,
XX comprises binding a molecule from library to a molecule having binding
XX property corresponding to CCR5 and identifying bound molecule -
XX
XX Example 3; Fig 4B; 50pp; English.
XX
XX The invention relates to a method for identifying a binding compound
XX for CC chemokine receptor 5 (CCR5). The method involves screening a
XX library of test molecules (particularly peptides) with immobilised CCR5,
XX and then identifying those molecules which bind. The invention also
XX relates to CCR5-binding molecules identified using the method of the
XX invention, methods for identifying consensus motifs for CCR5-binding
XX peptides, a transfer vector encoding tagged CCR5, a computer-aided
XX methods for determining the relative binding affinity of a test molecule
XX to CCR5 and a computer aided drug screening assay that utilises the
XX three-dimensional structure of CCR5. Compounds identified using the
XX methods of the invention are useful for treating or preventing HIV
XX (human immunodeficiency virus) infection or AIDS (acquired
XX immunodeficiency syndrome) in a patient. The methods of the invention
XX may also be used to identify agonists or antagonists of the interaction
XX of CCR5 with its natural ligand, and to determine a binding motif for
XX CCR5. The present sequence represents a naturally occurring variant of
XX human CCR5 in which there is a glutamine, rather than a leucine, at
XX position 55.
XX
XX SQ Sequence 352 AA;

Alignment Scores:

| | | | |
|------------------------|----------|---------------|-----|
| Pred. No.: | 1.6e-186 | Length: | 352 |
| Score: | 1739.50 | Matches: | 339 |
| Percent Similarity: | 96.32% | Conservative: | 1 |
| Best Local Similarity: | 96.03% | Mismatches: | 1 |
| Query Match: | 66.19% | Indels: | 12 |
| DB: | 23 | Gaps: | 1 |

US-09-938-703-3 (1-1442) x AAM52829 (1-352)

QY 240 ATGGATTATCAAGTGTCAAGTCCAAATCTATGACATCAATATTATACATCGAGCCCTGC 299
DB 1 MetAspTyrGlnValSerSerProIleTyrAspIleAsnTyrTyrThrSerGluProCys 20
QY 300 CAAAAATCAATGTAAGCAAAATCGAGCCGCCCTCCCTGCTCGCTCTACTACTGGTG 359
DB 21 GlnLysIleAsnValGlnIleAlaAlaArgLeuLeuProLeuTyrSerLeuVal 40
QY 360 TTCATCTTTGGTTTGTGGCAACATGCTGGTGCATCTCCATCTCTGATGAACGCAAGG 419
DB 41 PheIlePheGlyPheValGlyAsnMetLeuValIleLeuIleGlnIleAsnCysLysArg 60
QY 420 CTGAGAGCATGACATCTACTCTCTCAACCTGGCCATCTCTGACCTGTTTCTTCTT 479
DB 61 LeuLysSerMetThrAspIleTyrLeuLeuAsnLeuAlaIleSerAspLeuPheLeu 80
QY 480 CTTACTGTCCCTTCTGGGCTCACTATGCTGGCGCCAGTGGGACTTTGGAATACATG 539
DB 81 LeuThrValProPheThrPalaHisTyrAlaAlaAlaGlnTrpAspPheGlyAsnThrMet 100
QY 540 TCTCAACTCTTGACAGGCTCTATTTATAGGCTTCTTCTCTGGAATCTTCTCATCATC 599
DB 101 CysGlnLeuLeuThrGlyLeuTyrPheIleGlyPheSerGlyIlePhePheIleIle 120
QY 600 CTCCTGACAATCGATAGGTACCTGGCTGCTCATGCTGTGTGTTGTTTAAAGCCAGG 659
DB 121 LeuLeuThrIleAspArgTyrLeuAlaValValHisAlaValPheAlaLeuLysAlaArg 140
QY 660 ACGGTCACCTTTGGGCTGGTGCACAAAGTGTGATCCTGGTGGTGGTGGTGGTGGTCT 719
DB 141 ThrValThrPheGlyValValThrSerValIleThrTrpValValAlaValPheAlaSer 160
QY 720 CTCOCAGGAATCATCTTTACAGATCTCAAAAGAGAGGTCTTATTACACTGCACCTCT 779
DB 161 LeuProGlyIleIlePheThrArgSerGlnLysGluGlyLeuHisTyrThrCysSerSer 180
QY 780 CATTTCCTCATAC-----ATTAAAGATAGTCAT 806
DB 181 HisPheProTyrSerGlnTyrGlnPheThrPheTrpLysAsnPheGlnThrLeuLysIleVal 200
QY 807 CTTGGGGCTGGTCTCGCTGCTTGTTCATGCTCATCTCTACTCTCGGGAATCCTAAAC 866
DB 200 eLeuGlyLeuValLeuProLeuLeuValMetValIleCysTyrSerGlyIleLeuLys 220
QY 867 TCTGCTTCGGTGTGCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 926
DB 220 rLeuLeuArgCysArgAsnGlnLysLysArgHisArgAlaValArgLeuIlePheThrIle 240
QY 927 CATGATTGTTTATTTCTCTTCTGGGCTCCCTCAACATGCTCTCTCTCTGCAACCTT 986
DB 240 eMetIleValTyrPheLeuPheIleTrpAlaProTyrAsnIleValLeuLeuLeuAsnThrPh 260
QY 987 CCAGGAATTTCTTGGCTGAAATTAATTCAGTAGCTCTACAGGTTGGACCAAGCTATGCA 1046
DB 260 eGlnGluPhePheGlyLeuAsnAsnCysSerSerSerAsnArgLeuAspGlnAlaMetG 280
QY 1047 GGTTCAGAGAGATCTTGGATGAGCAGTGTGATCAACCCCATCATCTATGCTTTGT 1106
DB 280 nValThrGluThrLeuGlyMetThrHisCysCysIleAsnProIleIleTyrAlaPheVa 300
QY 1107 CGGGAGAAAGTTTCAGAAACTACCTCTTAGTCTTCTTCCAAAAGACATGCAACAGCTT 1166
DB 300 lGlyGluLysPheArgAsnTyrLeuLeuValPhePheGlnLysHisIleAlaLysArgPh 320
QY 1167 CTGCAAAATGCTTCTTATTTTCCAGAGAGGCTCCCGAGGAGCAAGCTCAGTTTACAC 1226
DB 320 eCysLysCysCysSerIlePheGlnGlnAlaProGluArgAlaSerSerValTyrTh 340
QY 1227 CCGATCCACTGGGAGCAGGAAATATCTGTGGGCTG 1263
DB 340 rArgSerThrGlyGluGlnGluIleSerValGlyLeu 352

RESULT 15

AAE07039

ID AAE07039 standard; Protein; 352 AA.

XX AC

XX AAE07039;

XX DT

XX 16-OCT-2001 (first entry)

XX DE

XX Human G-protein chemokine receptor (CCR5) HDGNR10 protein #2.

XX KW

Human; G-protein chemokine receptor; CCR5; HDGNR10; inflammation; HIV; human immunodeficiency virus; antimicrobial; vasodilator; vulnerability; cytotactic; immunosuppressive; nontropic; neuroprotective; gene therapy; neurodegenerative disorder; Kaposi's sarcoma; autoimmune disease; rheumatoid arthritis; cancer; breast; ovary; adrenal gland; bone marrow; gastrointestinal tract; lung; liver; immune disorder; Addison's disease; haemolytic anaemia; autoimmune thyroiditis; diabetes mellitus; allergy; multiple sclerosis; ulcerative colitis; Crohn's disease; wound healing; cardiovascular disorder; myocardial ischaemia.

XX OS

XX Homo sapiens.

XX PW

XX WO200158915-A2.

XX PD

XX 16-AUG-2001.

XX PF

XX 09-FEB-2001; 2001WO-US04152.

XX PR

XX 09-FEB-2000; 2000US-0181258.

XX PR

XX 09-MAR-2000; 2000US-0187999.

XX PR

XX 22-SEP-2000; 2000US-0234336.

XX PA

XX (HUMA-) HUMAN GENOME SCI INC.

XX PI

XX Rosen CA, Roschke V, Li Y, Ruben SM;

XX DR

XX WPI; 2001-489865/53.

XX DR

XX N-PSDB; AAD13198.

XX XX

XX Isolated nucleic acid encoding a human G-protein chemokine receptor

XX PT

XX (CCR5) HDGNR10 polypeptide, useful for preventing or treating

XX PT

XX autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative

XX PT

XX disorders and neurodegenerative disorders

XX PS

XX Example 40; Page 486-487; 495pp; English.

XX CC

XX The present sequence is human G-protein chemokine receptor (CCR5) HDGNR10

XX CC

XX protein. CCR5 HDGNR10 antibodies are useful for treating, preventing

XX CC

XX or ameliorating a disease or disorder associated with inflammation,

XX CC

XX defective or aberrant chemotaxis of immune cells, HIV infection (such as

XX CC

XX Pneumocystis carinii pneumonia or Kaposi's sarcoma) or defective or

XX CC

XX aberrant T-cell antigen presenting cell interaction. The disease or

XX CC

XX disorder may also be an infectious disease (e.g. a viral infection such

XX CC

XX as an early stage HIV infection, a cytomegalovirus infection, or a

XX CC

XX poxvirus infection) an autoimmune disease (e.g. rheumatoid arthritis) or

XX CC

XX a neurodegenerative disorder. The disease or disorder may be associated

XX CC

XX with aberrant CCR5 expression, lack of CCR5 function, aberrant CCR5

XX CC

XX ligand expression, or lack of CCR5 ligand function. CCR5 HDGNR10 protein

Alignment Scores:

Pred. No.: 2,69e-186 Length: 352
Score: 1737.50 Matches: 339
Percent Similarity: 96.32% Conservative: 1
Best Local Similarity: 96.03% Mismatches: 1
Query Match: 66.11% Indels: 12
DB: 22 Gaps: 1

US-09-938-703-3 (1-1442) x AAE07039 (1-352)

QY 240 ATGGATTATCAAGTGTCAAGTCCAACTATGACATCAATATTATACATGAGGCCCTGC 299
DB 1 MetAspTyrGlnValSerSerProIleTyrAspIleAsnTyrThrSerGluProCys 20
QY 300 CAAATAATCAATGTGAGCAATATCCAGCCGGCTCTCCCTCGCTCTACTCATGTGG 359
DB 21 GlnLysIleAsnValLysGlnIleAlaAlaArgLeuProLeuTyrSerLeuVal 40
QY 360 TTTCATCTTTGGTTTGGGCAACATGCTGTCATCTCATCTCATCTGATAAACTCAAAAGG 419
DB 41 PheIlePheGlyPheValGlyAsnMetLeuValIleLeuIleAsnCysLysArg 60
QY 420 CTGAAGAGCATGACTGACATCTACCTGCTCAACCTGGCCATCTCTGACCTGTTTCCTT 479
DB 61 LeuLysSerMetThrAspIleTyrLeuLeuAsnLeuAlaIleSerAspLeuPheLeu 80
QY 480 CTTACTGTCCCTTCTGGGCTCACTATGCTGCCGCCAGTGGACTTTGGAAATACATG 539
DB 81 LeuThrValProPheTrpAlaHisTyrAlaAlaIleTrpAspPheGlyAsnThrMet 100
QY 540 TGCAACTCTTGACAGGGCTCTATTTATAGGCTTTCTCTGGAATCTCTTCATCATC 599
DB 101 CysGlnLeuLeuThrGlyLeuTyrPheIleGlyPheSerGlyPhePheIleIle 120
QY 600 CTCCTGACAATCGATAGGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 659
DB 121 LeuLeuThrIleAspArgTyrLeuAlaValValHisAlaValPheAlaLeuLysAlaArg 140
QY 660 ACGSTCACTTTGGGTGGTGACAGTGTGATCACTTGGGTGGTGGTGGTGGTGGTGGT 719
DB 141 ThrValThrPheGlyValValThrSerValIleThrTrpValValAlaValPheAlaSer 160
QY 720 CTCCTGAGGAATCATCTTTACAGATCTCAAAAGAGGTCTTCATTCACCTGAGCTCT 779
DB 161 LeuProGlyIleIlePheThrArgSerGlnLysGlyLeuHisTyrThrCysSerSer 180
QY 780 CATTTTCCATAC-----ATTAAAGATGTCAT 806
DB 181 HisPheProTyrSerGlnTyrGlnPheTrpLysAsnPheGlnThrLeu-LysIleVal 200
QY 807 CTTGGGCTGGTCTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 866
DB 200 eLeuGlyLeuValLeuProLeuLeuValMetValIleCysTyrSerGlyIleLeuLys 220
QY 867 TCTGCTTCGGTGTGAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 926
DB 220 rLeuLeuArgCysArgAsnGlnLysLysArgHisArgAlaValArgLeuIlePheThr 240
QY 927 CATGATTTGTTATTTCTCTGCTGGCTCCCTACACATGCTCTCTCTCTGCAACCTT 986
DB 240 eMetIleValTyrPheLeuPheTrpAlaProTyrAsnIleValLeuLeuLeuAsnThrPh 260
QY 987 CCAGGAATCTTTGGCTGAAATATTCAGTAGCTCTAACAGGTTGGACCAAGCTATGCA 1046
DB 260 eGlnGluPhePheGlyLeuAsnAsnCysSerSerSerAsnArgLeuAspGlnAlaMetG 280
QY 1047 GTCACAGAGACTCTGGGATGAGGACGACGACGACGACGACGACGACGACGACGACG 1106
DB 280 nValThrGluThrLeuGlyMetThrHisCysCysIleAsnProIleIleTyrAlaPheVa 300
QY 1107 CGGGGAGAGTTCAGAACTACTCTTACTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1166
DB 300 lGlyGluLysPheArgAsnTyrLeuLeuValPhePheGlnLysHisIleAlaLysArgPh 320

| | | | |
|----|------|--|------|
| QY | 1167 | CTGCAATGCTGTTCTATTTCAGCAAGAGGCTCCGAGCGAGCAAGCTCAGTTTACAC | 1226 |
| Db | 320 | eCysLysCysCysSerIlePheGlnGlnGluAlaProGluArgAlaSerSerValTyrTh | 340 |
| QY | 1227 | CCGATCCACTGGGGAGCAGGAATATCTGTGGGCTTG | 1263 |
| Db | 340 | rArgSerThrGlnGlnGlnIleSerValGlyLeu | 352 |

Search completed: June 3, 2003, 18:45:05
Job time : 91.6348 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_p2p model.

Run on: June 3, 2003, 18:41:24 ; Search time 6.73177 Seconds
(without alignments)
5754.334 Million cell updates/sec

Title: US-09-938-703-3_COPY_792_885

Perfect score: 164

Sequence: 1 ATTAAGATAGTCACTCTGG.....CTCTGCTTCGGTGCAGAAAT 94

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 1343160

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+np2p.model -DEV=xlp
-O=/cgn2_1/USPTO_spool/US09938703/runat_03062003_161407_22827/app_query.fasta.1.1870
-DB=SPREMBL_21 -OPMT=fastan -SUFFIX=rspt -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09938703.ecgn.1.1.192.etrnat_03062003_161407_22827 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6
-Fgapext=7 -Ygapop=10 -Ygapext=0.5 -DELOP=6 -DELEXT=7

Database :

SPREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Query | Score | Match | Length | ID | Description |
|------------|-------|-------|-------|--------|--------|--------------------|
| 1 | 164 | 100.0 | 215 | 4 | 075303 | 075303 homo sapien |

| | | | | | | |
|----|-----|------|-----|---|--------|--------------------|
| 2 | 152 | 92.7 | 60 | 4 | 09PLT4 | 09PLT4 homo sapien |
| 3 | 152 | 92.7 | 333 | 4 | 014694 | 014694 homo sapien |
| 4 | 152 | 92.7 | 339 | 4 | Q9UN28 | Q9UN28 homo sapien |
| 5 | 152 | 92.7 | 339 | 4 | Q9UN27 | Q9UN27 homo sapien |
| 6 | 152 | 92.7 | 339 | 4 | Q9UN26 | Q9UN26 homo sapien |
| 7 | 152 | 92.7 | 339 | 4 | Q9UN25 | Q9UN25 homo sapien |
| 8 | 152 | 92.7 | 339 | 4 | Q9UN24 | Q9UN24 homo sapien |
| 9 | 152 | 92.7 | 339 | 4 | Q9UN23 | Q9UN23 homo sapien |
| 10 | 152 | 92.7 | 339 | 4 | Q9UBT9 | Q9UBT9 homo sapien |
| 11 | 152 | 92.7 | 339 | 4 | Q9UBJ7 | Q9UBJ7 homo sapien |
| 12 | 152 | 92.7 | 339 | 6 | Q9TUX1 | Q9TUX1 hylobates c |
| 13 | 152 | 92.7 | 339 | 6 | Q9TUX0 | Q9TUX0 hylobates c |
| 14 | 152 | 92.7 | 339 | 6 | Q9TUX9 | Q9TUX9 hylobates c |
| 15 | 152 | 92.7 | 339 | 6 | Q9TUM7 | Q9TUM7 pan troglod |
| 16 | 152 | 92.7 | 339 | 6 | Q9TUM6 | Q9TUM6 pan troglod |
| 17 | 152 | 92.7 | 339 | 6 | Q9TUM5 | Q9TUM5 pan troglod |
| 18 | 152 | 92.7 | 339 | 6 | Q9TUM4 | Q9TUM4 pongo pygma |
| 19 | 152 | 92.7 | 339 | 6 | Q9TUM3 | Q9TUM3 cercopithec |
| 20 | 152 | 92.7 | 339 | 6 | Q9TUM9 | Q9TUM9 cercopithec |
| 21 | 152 | 92.7 | 339 | 6 | Q9TUM8 | Q9TUM8 cercopithec |
| 22 | 152 | 92.7 | 339 | 6 | Q9TUM9 | Q9TUM9 papio cynoc |
| 23 | 152 | 92.7 | 339 | 6 | Q9TUM8 | Q9TUM8 papio cynoc |
| 24 | 152 | 92.7 | 339 | 6 | Q9TUM7 | Q9TUM7 papio cynoc |
| 25 | 152 | 92.7 | 339 | 6 | Q9TUM6 | Q9TUM6 papio cynoc |
| 26 | 152 | 92.7 | 339 | 6 | Q9TUM5 | Q9TUM5 papio cynoc |
| 27 | 152 | 92.7 | 339 | 6 | Q9TUM4 | Q9TUM4 cercopithec |
| 28 | 152 | 92.7 | 339 | 6 | Q9TUM3 | Q9TUM3 cercopithec |
| 29 | 152 | 92.7 | 339 | 6 | Q9TUM8 | Q9TUM8 cercopithec |
| 30 | 152 | 92.7 | 339 | 6 | Q9TUM1 | Q9TUM1 cercopithec |
| 31 | 152 | 92.7 | 339 | 6 | Q9TUM7 | Q9TUM7 cercopithec |
| 32 | 152 | 92.7 | 339 | 6 | Q9TUM6 | Q9TUM6 cercopithec |
| 33 | 152 | 92.7 | 339 | 6 | Q9TUM4 | Q9TUM4 mandrillus |
| 34 | 152 | 92.7 | 339 | 6 | Q9TUM3 | Q9TUM3 erythrocebu |
| 35 | 152 | 92.7 | 339 | 6 | Q9TUM2 | Q9TUM2 erythrocebu |
| 36 | 152 | 92.7 | 339 | 6 | Q9TUM9 | Q9TUM9 cercopithec |
| 37 | 152 | 92.7 | 339 | 6 | Q9TUM8 | Q9TUM8 cercopithec |
| 38 | 152 | 92.7 | 339 | 6 | Q9TUM5 | Q9TUM5 erythrocebu |
| 39 | 152 | 92.7 | 339 | 6 | Q9TUM6 | Q9TUM6 erythrocebu |
| 40 | 152 | 92.7 | 339 | 6 | Q9TUM4 | Q9TUM4 erythrocebu |
| 41 | 152 | 92.7 | 339 | 6 | Q9TUM3 | Q9TUM3 mandrillus |
| 42 | 152 | 92.7 | 339 | 6 | Q9TUM2 | Q9TUM2 erythrocebu |
| 43 | 152 | 92.7 | 339 | 6 | Q9TUM1 | Q9TUM1 pan troglod |
| 44 | 152 | 92.7 | 339 | 6 | Q9TUM4 | Q9TUM4 pongo pygma |
| 45 | 152 | 92.7 | 339 | 6 | Q9TUM2 | Q9TUM2 pongo pygma |

ALIGNMENTS

RESULT 1
075303
ID 075303 PRELIMINARY; PRT; 215 AA.
AC 075303;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE CC-chemokine receptor.
GN CCR-5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Tse L., Ehrenberg P.K., Chang G., Michael N.L.;
RT Genomic Organization and Functional Characterization of the Complete
RT Transcription Unit for the Chemokine Receptor CCR-5, a Major Entry Co-
RL Receptor for HIV-1.
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF009962; AAC23944.1; -;
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam: PF00001; 7tm1; 1.
DR PROSITE: PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE: PS02662; G_PROTEIN_RECEP_F1_2; 1.
DR

| | | | | |
|----|----------|---------|-----------|-------------------------|
| FT | NON_TER | 1 | 1 | |
| FT | NON_TER | 339 | 339 | 55C819F92D6DB1A6 CRC64: |
| SO | SEQUENCE | 339 AA: | 30183 MW: | |

Alignment Scores:
 Pred. No.: 3e-14 Length: 339
 Score: 152.00 Matches: 31
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 92.68% Indels: 0
 DB: 4 Gaps: 0

US-09-938-703-3_COPY_792_885 (1-94) x Q9UN28 (1-339)

QY 2 TTAAGATAGTCATCTGGGCGTGCCTGGCTGCTCATGTCATCTGCTACTCG 61

DB 189 LeuylsIlleValleuGlyLeuValleuProleuValMetValleuCysTyrSer 208

QY 62 GGAATCCCTAAACCTCTGCTGCTGTCGAAAT 94

DB 209 GlylleLeuLysThrLeuLeuArgCysArgAsn 219

RESULT 5

Q9UN27
 ID Q9UN27 PRELIMINARY; PRT; 339 AA.
 AC Q9UN27;
 DT 01-MAY-2000 (TReMBLrel. 13, Created)
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
 DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
 DE C-C chemokine receptor 5 (Fragment).
 GN CCR5.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]

RP SEQUENCE FROM N.A.
 RA Kunstman K., Chen Z., Korber B., Oprendeck J., Stanton J., Agy M.,
 RA Shibata R., Yoder A., Pillai S., Kuiken C., Marx P., Wolinsky S.;
 RT "Sequences of the CCR5 genes from diverse simian and prosimian
 RT species.";
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF161914; AAD47671.1;
 DR InterPro; IPR000276; GPCR_Rhodpsn.
 DR PRINTS; PS00237; GPCRHHODPSN.
 DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
 DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
 KW Receptor.
 FT NON_TER 1
 FT NON_TER 339
 SQ SEQUENCE 339 AA; 39086 MW; 88AD8B44E2CB4EC2 CRC64;

Alignment Scores:
 Pred. No.: 3e-14 Length: 339
 Score: 152.00 Matches: 31
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 92.68% Indels: 0
 DB: 4 Gaps: 0

US-09-938-703-3_COPY_792_885 (1-94) x Q9UN27 (1-339)
 QY 2 TTAAGATAGTCATCTGGGCGTGCCTGGCTGCTCATGTCATCTGCTACTCG 61

DB 189 LeuylsIlleValleuGlyLeuValleuProleuValMetValleuCysTyrSer 208

QY 62 GGAATCCCTAAACCTCTGCTGCTGTCGAAAT 94

DB 209 GlylleLeuLysThrLeuLeuArgCysArgAsn 219

RESULT 6

Q9UN26
 ID Q9UN26 PRELIMINARY; PRT; 339 AA.
 AC Q9UN26;
 DT 01-MAY-2000 (TReMBLrel. 13, Created)

DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
 DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
 DE C-C chemokine receptor 5 (Fragment).
 GN CCR5.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]

RP SEQUENCE FROM N.A.
 RA Kunstman K., Chen Z., Korber B., Oprendeck J., Stanton J., Agy M.,
 RA Shibata R., Yoder A., Pillai S., Kuiken C., Marx P., Wolinsky S.;
 RT "Sequences of the CCR5 genes from diverse simian and prosimian
 RT species.";
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF161916; AAD47673.1;
 DR InterPro; IPR000276; GPCR_Rhodpsn.
 DR Pfam; PF00001; 7tm_1; 1.
 DR PRINTS; PS00237; GPCRHHODPSN.
 DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
 DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
 KW Receptor.
 FT NON_TER 1
 FT NON_TER 339
 SQ SEQUENCE 339 AA; 39162 MW; A56369FE0529F4AB CRC64;

Alignment Scores:
 Pred. No.: 3e-14 Length: 339
 Score: 152.00 Matches: 31
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 92.68% Indels: 0
 DB: 4 Gaps: 0

US-09-938-703-3_COPY_792_885 (1-94) x Q9UN26 (1-339)

QY 2 TTAAGATAGTCATCTGGGCGTGCCTGGCTGCTCATGTCATCTGCTACTCG 61

DB 189 LeuylsIlleValleuGlyLeuValleuProleuValMetValleuCysTyrSer 208

QY 62 GGAATCCCTAAACCTCTGCTGCTGTCGAAAT 94

DB 209 GlylleLeuLysThrLeuLeuArgCysArgAsn 219

RESULT 7

Q9UN25
 ID Q9UN25 PRELIMINARY; PRT; 339 AA.
 AC Q9UN25;
 DT 01-MAY-2000 (TReMBLrel. 13, Created)
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
 DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
 DE C-C chemokine receptor 5 (Fragment).
 GN CCR5.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]

RP SEQUENCE FROM N.A.

RA Kunstman K., Chen Z., Korber B., Oprendeck J., Stanton J., Agy M.,
 RA Shibata R., Yoder A., Pillai S., Kuiken C., Marx P., Wolinsky S.;
 RT "Sequences of the CCR5 genes from diverse simian and prosimian
 RT species.";
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF161918; AAD47675.1;
 DR InterPro; IPR000276; GPCR_Rhodpsn.
 DR Pfam; PF00001; 7tm_1; 1.
 DR PRINTS; PS00237; GPCRHHODPSN.
 DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
 DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
 KW Receptor.
 FT NON_TER 1
 FT NON_TER 339

| | |
|-----------|---------------------------|
| RESULT 5 | |
| Q9UN23 | |
| ID Q9UN23 | PRELIMINARY; PRT; 339 AA. |
| AC Q9UN23 | |

```
DR pfam; PF00001; 7tm_1; 1.  
DR PRINTS; PR00327; CPOBHHOBSN
```

RESULT 9
Q9UN23
ID Q9UN23
AC Q9UN23.

100

SEQUENCE FROM N.A.
RA Kunstan K., Chen Z., Korber B., Oprondek J., Stanton J., Agly M.,
RA Shibata R., Yoder A., Pillai S., Kuiken C., Marx P., Wolinsky S.;
RT "Sequences of the CCR5 genes from diverse simian and prosimian
RT species";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF161888; AADA7645.1; -
DR InterPro: IPR000276; GPCR_Rhodpsn.

Search completed: June 3, 2003, 18:49:49
Job time : 7.73177 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus.p2p model

Run on: June 3, 2003, 18:39:54 ; Search time 1.52995 Seconds
(without alignments)
5096.611 Million cell updates/sec

Title: US-09-938-703-3_COPY_792_885

Perfect score: 164

Sequence: 1 ATTAAGATAGTATCTTGG.....CTCTGCTTCGTGTCGAAT 94

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 225784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame-n2p.model -DEV=xlp
-O=/cgm2_1/USPTO.spool/US09938703/runat_03062003_161406_22808/app_query.fasta_1.1870
-DB=SwissProt_40 -QFMT=Eastan -SUFFIX=rs -MINMATCH=0.1 -LOOPPC1=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCLALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09938703 -CGN_1_1_38 -runat_03062003_161406_22808 -NCPO=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARM_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------|
| 1 | 152 | 92.7 | 352 | 1 | CCR5_CERAE |
| 2 | 152 | 92.7 | 352 | 1 | CCR5_CERTO |
| 3 | 152 | 92.7 | 352 | 1 | CCR5_GORGO |
| 4 | 152 | 92.7 | 352 | 1 | CCR5_HUMAN |
| 5 | 152 | 92.7 | 352 | 1 | CCR5_HYLLE |
| 6 | 152 | 92.7 | 352 | 1 | CCR5_PANTR |
| 7 | 152 | 92.7 | 352 | 1 | CCR5_PAPHA |
| 8 | 152 | 92.7 | 352 | 1 | CCR5_PONPY |
| 9 | 152 | 92.7 | 352 | 1 | CCR5_PYGBI |
| 10 | 152 | 92.7 | 352 | 1 | CCR5_TRAFH |
| 11 | 152 | 92.7 | 352 | 1 | CCR5_TRAFH |
| 12 | 151 | 92.1 | 352 | 1 | CCR5_PYGNE |
| 13 | 149 | 90.9 | 352 | 1 | CCR5_PYGNE |
| 14 | 134 | 81.7 | 360 | 1 | CCR2_MACMU |
| 15 | 134 | 81.7 | 374 | 1 | CCR2_HUMAN |
| 16 | 133 | 81.1 | 354 | 1 | CCR5_RAT |
| 17 | 131 | 79.9 | 354 | 1 | CCR5_MOUSE |
| 18 | 122 | 74.4 | 359 | 1 | CCR5_MOUSE |

| | | | | | |
|----|-----|------|-----|---|------------|
| 19 | 121 | 73.8 | 359 | 1 | CCR3_RAT |
| 20 | 118 | 72.0 | 373 | 1 | CCR2_MOUSE |
| 21 | 118 | 72.0 | 373 | 1 | CCR2_RAT |
| 22 | 113 | 68.9 | 358 | 1 | CCR3_CAVPO |
| 23 | 112 | 68.3 | 355 | 1 | CCR3_MACMU |
| 24 | 109 | 66.5 | 355 | 1 | CCR3_CERAE |
| 25 | 109 | 66.5 | 355 | 1 | CCR3_HUMAN |
| 26 | 106 | 64.6 | 355 | 1 | CCR1_HUMAN |
| 27 | 105 | 64.0 | 355 | 1 | CCR1_MACMU |
| 28 | 105 | 64.0 | 355 | 1 | CCR1_MOUSE |
| 29 | 95 | 57.9 | 360 | 1 | CCR4_HUMAN |
| 30 | 88 | 53.7 | 354 | 1 | C3X1_RAT |
| 31 | 84 | 51.2 | 354 | 1 | C3X1_MOUSE |
| 32 | 81 | 49.4 | 356 | 1 | CKEV_MOUSE |
| 33 | 80 | 48.8 | 342 | 1 | CCR6_CERAE |
| 34 | 80 | 48.8 | 342 | 1 | CCR6_HUMAN |
| 35 | 80 | 48.8 | 342 | 1 | CCR6_MACNE |
| 36 | 80 | 48.8 | 343 | 1 | CCR6_MACMU |
| 37 | 80 | 48.8 | 360 | 1 | CCR4_MOUSE |
| 38 | 79 | 48.2 | 269 | 1 | VC03_SPVKA |
| 39 | 79 | 48.2 | 355 | 1 | C3X1_HUMAN |
| 40 | 78 | 47.6 | 370 | 1 | VK02_SPVKA |
| 41 | 76 | 46.3 | 378 | 1 | CCR7_HUMAN |
| 42 | 75 | 45.7 | 355 | 1 | CCR8_HUMAN |
| 43 | 74 | 45.1 | 374 | 1 | CCR6_HUMAN |
| 44 | 72 | 43.9 | 356 | 1 | CCR8_MACMU |
| 45 | 71 | 43.3 | 378 | 1 | CCR7_MOUSE |

ALIGNMENTS

RESULT 1
ID CCR5_CERAE STANDARD; PRT; 352 AA.
AC P56493;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE C-C chemokine receptor type 5 (C-C CCR-5) (CCR-5) (CCR5).
GN CCR5 OR CCR5R5.
OS Cercopithecus aethiops (Green monkey) (Grivet).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC Cercopithecinae; Cercopithecus.
OX NCBI_TaxID=9534;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=98001387; PubMed=9343222;
RA Kuhlmann S.E., Platt E.J., Kozak S.L., Kabat D.;
RT "Polymorphisms in the CCR5 genes of African green monkeys and mice
implicate specific amino acids in infections by simian and human
immunodeficiency viruses.";
RL J. Virol. 71:8642-8656(1997).
RN [2]

SEQUENCE FROM N.A.
RP Murayama Y., Matsunaga S., Inoue-Murayama M.;
RA "cDNA sequence of African green monkey CCR-5 chemokine receptor
gene.";
RT Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: RECEPTOR FOR A C-C TYPE CHEMOKINE. BINDS TO MIP-1-ALPHA,
MIP-1-BETA AND RANTES AND SUBSEQUENTLY TRANSDUCES A SIGNAL BY
INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL. MAY PLAY A ROLE
IN THE CONTROL OF GRANULOCYTIC LINEAGE PROLIFERATION OR
DIFFERENTIATION.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial

QY 62 GGAATCCTAAACACTCTGCTCGTGCAGAAAT 94
 |||||||
 Db 216 GlyIleLeuLysThrLeuLeuArgCysArgAsn 226

RESULT 3
 CKR5_GORGO STANDARD; PRT; 352 AA.
 ID CKR5_HUMAN
 AC P56439;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 13-JUL-1998 (Rel. 36, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE C-C chemokine receptor type 5 (CCR-5) (CCR-5) (CCR5).
 GN CCR5 OR CMKBR5.
 OS Gorilla gorilla gorilla (Lowland gorilla).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
 OX NCBI_TaxID=9595;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97268687; PubMed=9108095;
 RA Edinger A.L., Amedee A., Miller K., Doranz B.J., Endres M.,
 RA Sharper M., Samson M., Lu Z.H., Clements J.E., Murphy-Corb M.,
 RA Pelper S.C., Parmentier M., Broder C.C., Doms R.W.;
 RT "Differential utilization of CCR5 by macrophage and T cell tropic
 RT simian immunodeficiency virus strains.";
 RL Proc. Natl. Acad. Sci. U.S.A. 94:4005-4010(1997).
 CC MIP-1-BETA AND RANTES AND SUBSEQUENTLY TRANSDUCE A SIGNAL BY
 CC INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL. MAY PLAY A ROLE
 CC IN THE CONTROL OF GRANULOCYTIC LINEAGE PROLIFERATION OR
 CC DIFFERENTIATION.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF005659; AAB62553.1;
 CC InterPro; IPR000276; GPCR_Rhodopsn.
 CC Pfam; PF00001; 7tm_1; 1.
 CC PRINTS; PRO0237; GPCR_Rhodopsn.
 CC PROSITE; PS00237; G-PROTEIN_RECEP_F1_1; 1.
 CC PROSITE; PS00262; G-PROTEIN_RECEP_F1_2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation.
 FT DOMAIN 1 30
 FT TRANSMEM 31 58
 FT DOMAIN 59 68
 FT TRANSMEM 69 89
 FT DOMAIN 90 102
 FT TRANSMEM 103 124
 FT DOMAIN 125 141
 FT TRANSMEM 142 166
 FT DOMAIN 167 198
 FT TRANSMEM 199 218
 FT DOMAIN 219 235
 FT TRANSMEM 236 260
 FT DOMAIN 261 277
 FT TRANSMEM 278 301
 FT DOMAIN 302 352
 FT DISULFID 101 178
 FT MOD_RES 3 3
 FT MOD_RES 10 10
 FT MOD_RES 14 14
 FT MOD_RES 15 15
 SQ SEQUENCE 352 AA; 40515 MW; D0B6FCB9FE5EAC84 CRC64;

Alignment Scores:

Pred. No.: 5 57e-14 Length: 352
 Score: 152.00 Matches: 31
 Percent Similarity: 100.00% Conservatives: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 92.68% Indels: 0
 DB: 1 Gaps: 0

US-09-938-703-3_COPY_792_885 (1-94) x CKR5_GORGO (1-352)

QY 2 TTAAAGATAGTCATCTTGGGCTGCTCGTCCCTGCTGTCATGTCATCTGCTACTCG 61
 |||||||
 Db 196 LeuLysIleValIleLeuGlyLeuValLeuProLeuLeuValMetValIleCysTyrSer 215

QY 62 GGAATCCTAAACACTCTGCTCGTGCAGAAAT 94
 |||||||
 Db 216 GlyIleLeuLysThrLeuLeuArgCysArgAsn 226

RESULT 4
 CKR5_HUMAN STANDARD; PRT; 352 AA.
 ID CKR5_HUMAN
 AC P51681; O14692; O14693; O14695; O14696; O14697; O14698; O14699;
 AC O14700; O14701; O14702; O14703; O14704; O14705; O14706; O14707;
 AC O14708; O15538; Q9UPA4;
 DT 01-CCR-1996 (Rel. 34, Created)
 DT 01-CCR-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE C-C chemokine receptor type 5 (C-C CCR-5) (CCR-5) (CCR5)
 DE (HIV-1 fusion co-receptor) (CHEMR13) (CD195 antigen).
 GN CCR5 OR CMKBR5.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96241590; PubMed=8639485;
 RA Samson M., Labbe O., Mollereau C., Vassart G., Parmentier M.;
 RT "Molecular cloning and functional expression of a novel human
 RT C-C chemokine receptor gene.";
 RL Biochemistry 35:3362-3367(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96291862; PubMed=8663314;
 RA Raport C.J., Gosling J., Schweichart V.L., Gray P.W., Charo I.F.;
 RT "Molecular cloning and functional characterization of a novel human
 RT C-C chemokine receptor (CCR5) for RANTES, MIP-1beta, and MIP-1alpha.";
 RL J. Biol. Chem. 271:17161-17166(1996).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96295970; PubMed=8699119;
 RA Combadiere C., Ahuja S.K., Tiffany H.L., Murphy P.M.;
 RT "Cloning and functional expression of CC CKR5, a human monocyte CC
 RT chemokine receptor selective for MIP-1(alpha), MIP-1(beta), and
 RT RANTES.";
 RL J. Leukoc. Biol. 60:147-152(1996).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX McCombie W.R., Wilson R., Chen E., Gibbs R., Zuo L., Johnson D.,
 RX Nham M., Farnell L., Dedhia N., Ansari A., Mardis E., Schutz K.,
 RX Guoj L., la Bastide M., Kaplan N., Greco T., Touchman J.,
 RX Muzny D., Chen C.N., Evans C., Fitzgerald M., See L.H., Tang M.,
 RX Sorci B.M., Dragan Y., Giacalone J., Pae A., Powell E.,
 RX Solinsky K.A., Desilva U., Diaz-Perez S., Zhou X., Yu Y.,
 RX Watanabe M., Doggett N., Garcia D., Sagripanti J.L.;
 RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98001387; PubMed=9343222;
 RX Kuhnann S.E., Platt E.J., Kozak S.L., Kabat D.;
 RT "Polymorphisms in the CCR5 genes of African green monkeys and mice
 RT implicate specific amino acids in infections by simian and human
 RT immunodeficiency viruses.";
 RL J. Virol. 71:8642-8656(1997).

| | |
|----|---|
| DR | EMBL; AF079431; AAU19863.1; |
| DR | InferPro; IPR000276; GPCR_Rhodpsd. |
| DR | Pfam; PF00001; 7tm_1.1. |
| DR | PRINTS; PR00237; GPCRRHOOPSN. |
| DR | PROSITE; PS00237; G_PROTEIN_RECP_FL_1. |
| DR | PROSITE; PS02362; G_PROTEIN_RECP_FL_2. |
| KW | G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation. |
| FT | DOMAIN 1 30 EXTRACELLULAR (POTENTIAL). |
| FT | TRANSMEM 31 58 1 (POTENTIAL). |
| FT | DOMAIN 59 68 CYTOPLASMIC (POTENTIAL). |
| FT | TRANSMEM 69 89 2 (POTENTIAL). |
| FT | DOMAIN 90 102 EXTRACELLULAR (POTENTIAL). |
| FT | TRANSMEM 103 124 3 (POTENTIAL). |
| FT | DOMAIN 125 141 CYTOPLASMIC (POTENTIAL). |
| FT | TRANSMEM 142 166 4 (POTENTIAL). |
| FT | DOMAIN 167 198 EXTRACELLULAR (POTENTIAL). |
| FT | TRANSMEM 199 218 5 (POTENTIAL). |
| FT | DOMAIN 219 235 CYTOPLASMIC (POTENTIAL). |
| FT | TRANSMEM 236 260 6 (POTENTIAL). |

RT "Chimpanzee CXCR4 and CCR5 act as coreceptors for HIV type 1." *N.*
RL AIDS Res. Hum. Retroviruses 13:1583-1587(1997).

CC -1- FUNCTION: RECEPTOR FOR A C-C TYPE CHEMOKINE. BINDS TO MIP-1-ALPHA,
 CC MIP-1-BETA AND RANTES AND SUBSEQUENTLY TRANSDUCES A SIGNAL BY
 CC INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL. MAY PLAY A ROLE
 CC IN THE CONTROL OF GRANULOCYTIC LINEAGE PROLIFERATION OR
 CC DIFFERENTIATION
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AF005663; AAB62557.1; -
 DR EMBL; U94329; AAB58446.1; -
 DR EMBL; AF011542; AAB65742.1; -
 DR EMBL; U97666; AAC51670.1; -
 DR EMBL; AF011540; AAB65740.1; -
 DR EMBL; U89797; AAC03717.1; -
 DR InterPro: IPR000276; GPCR_Rhodpsn.
 DR Pfam: PF00001; 7tm.1; 1.
 DR PRINTS: PR00237; GPCRHOOPS.
 DR PROSITE: PS00237; G-PROTEIN_RECEP_F1_1; 1.
 DR PROSITE: PS0262; G-PROTEIN_RECEP_F1_2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation.
 FT DOMAIN 1 30 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 31 58 1 (POTENTIAL).
 FT DOMAIN 59 68 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 69 89 2 (POTENTIAL).
 FT DOMAIN 90 102 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 103 124 3 (POTENTIAL).
 FT DOMAIN 125 141 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 142 166 4 (POTENTIAL).
 FT DOMAIN 167 198 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 199 218 5 (POTENTIAL).
 FT DOMAIN 219 235 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 236 260 6 (POTENTIAL).
 FT DOMAIN 261 277 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 278 301 7 (POTENTIAL).
 FT DOMAIN 302 352 CYTOPLASMIC (POTENTIAL).
 FT DISULFID 101 178 BY SIMILARITY.
 FT MOD_RES 3 3 SULFATION (BY SIMILARITY).
 FT MOD_RES 10 10 SULFATION (BY SIMILARITY).
 FT MOD_RES 14 14 SULFATION (BY SIMILARITY).
 FT MOD_RES 15 15 SULFATION (BY SIMILARITY).
 FT CARBOHYD 268 268 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 123 123 T -> S (IN REF. 1).
 SQ SEQUENCE 352 AA; 40539 MW; 4A33E698B80FE34C CRC64;

Alignment Scores:

Pred. No.: 5,57e-14 Length: 352
 Score: 152.00 Matches: 31
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 92.68% Indels: 0
 DB: 1 Gaps: 0

US-09-938-703-3_COPY_792_885 (1-94) x CKR5_PANTR (1-352)

QY 2 TTAAGATAGTCATCTGGGCTGGTCTCGCTGCTCATGCTACTGCTACTCG 61

Db 196 LeuylsIlleValIleLeuGlyLeuValLeuProLeuValMetValIleCysTyrSer 215

QY 62 GGAATCCCTAAACTCTGCTTGGTGCGAAT 94

Db 216 GlyIleLeuLysThrLeuLeuArgCysArgAsn 226

RESULT 7

CKR5_PAPHA STANDARD; PRT; 352 AA.
 ID CKR5_PAPHA

AC P56441;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE C-C chemokine receptor type 5 (C-CR-5) (CC-CCR-5) (CCR5-5) (CCR5).
 DE CCR5 OR CCR5R5.
 GN Papio hamadryas (Hamadryas baboon), and
 OS Papio anubis (Olive baboon).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecinae; Papio.
 OX NCBI_taxid-9557, 9555;
 RN [1]
 SEQUENCE FROM N.A.
 RP SPECIES-P.hamadryas;
 RC MEDLINE-97268687; PubMed-9108095;
 RA Edinger A.L., Amedee A., Miller K., Doranz B.J., Endres M.,
 RA Sharron M., Samson M., Lu Z.H., Clements J.E., Murphy-Corb M.,
 RA Reiser S.C., Parmentier M., Broder C.C., Doms R.W.;
 RT "Differential utilization of CCR5 by macrophage and T cell tropic
 RT simian immunodeficiency virus strains.";
 RL Proc. Natl. Acad. Sci. U.S.A. 94:4005-4010(1997).
 RN [2]
 SEQUENCE FROM N.A.
 RP SPECIES-P.hamadryas;
 RC MEDLINE-99210133; PubMed-10195758;
 RA Saksena N.K., Wang B., Novembre F.J., Bolton W., Smit T.K., Lal R.B.;
 RT "Species-specific changes in the CCR5 gene from African and Asian
 RT nonhuman primates.";
 RL AIDS Res. Hum. Retroviruses 15:479-483(1999).
 RN [3]
 SEQUENCE FROM N.A.
 RP SPECIES-P.anubis;
 RC MEDLINE-99210133; PubMed-10195758;
 RA Benton P.A., Timanus D.K., Shearer M.H., Lee D.R., Kennedy R.C.;
 RT Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: RECEPTOR FOR A C-C TYPE CHEMOKINE. BINDS TO MIP-1-ALPHA,
 CC MIP-1-BETA AND RANTES AND SUBSEQUENTLY TRANSDUCES A SIGNAL BY
 CC INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL. MAY PLAY A ROLE
 CC IN THE CONTROL OF GRANULOCYTIC LINEAGE PROLIFERATION OR
 CC DIFFERENTIATION.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

EMBL; AF005658; AAB62552.1; -
 EMBL; AF105287; AAD20556.1; -
 EMBL; AF105288; AAD20557.1; -
 EMBL; AF105289; AAD20558.1; -
 EMBL; AF105290; AAD20559.1; -
 EMBL; AF023452; AAC63830.1; -
 InterPro: IPR000276; GPCR_Rhodpsn.
 Pfam: PF00001; 7tm.1; 1.
 DR PRINTS: PR00237; GPCRHOOPS.
 DR PROSITE: PS00237; G-PROTEIN_RECEP_F1_1; 1.
 DR PROSITE: PS0262; G-PROTEIN_RECEP_F1_2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation.
 FT DOMAIN 1 30 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 31 58 1 (POTENTIAL).
 FT DOMAIN 59 68 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 69 89 2 (POTENTIAL).
 FT DOMAIN 90 102 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 103 124 3 (POTENTIAL).
 FT DOMAIN 125 141 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 142 166 4 (POTENTIAL).
 FT DOMAIN 167 198 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 199 218 5 (POTENTIAL).

| | | | | | |
|-----|----------|---------|-----------|------------------------|----------------------------|
| FT | DOMAIN | 1 | 30 | | EXTRACELLULAR (POTENTIAL). |
| FT | TRANSEM | 31 | 58 | | 1 (POTENTIAL). |
| FT | DOMAIN | 59 | 68 | | CYTOPLASMIC (POTENTIAL). |
| FT | TRANSEM | 69 | 89 | | 2 (POTENTIAL). |
| FT | DOMAIN | 90 | 102 | | EXTRACELLULAR (POTENTIAL). |
| FT | TRANSEM | 103 | 124 | | 3 (POTENTIAL). |
| FT | DOMAIN | 125 | 141 | | CYTOPLASMIC (POTENTIAL). |
| FT | TRANSEM | 142 | 166 | | 4 (POTENTIAL). |
| FT | DOMAIN | 167 | 198 | | EXTRACELLULAR (POTENTIAL). |
| FT | TRANSEM | 199 | 218 | | 5 (POTENTIAL). |
| FT | DOMAIN | 219 | 235 | | CYTOPLASMIC (POTENTIAL). |
| FT | TRANSEM | 236 | 260 | | 6 (POTENTIAL). |
| FT | DOMAIN | 261 | 277 | | EXTRACELLULAR (POTENTIAL). |
| FT | TRANSEM | 278 | 301 | | 7 (POTENTIAL). |
| FT | DOMAIN | 302 | 352 | | CYTOPLASMIC (POTENTIAL). |
| FT | DISULFID | 101 | 178 | | BY SIMILARITY. |
| FT | MOD_RES | 3 | 3 | | SULFATION (BY SIMILARITY). |
| FT | MOD_RES | 10 | 10 | | SULFATION (BY SIMILARITY). |
| FT | MOD_RES | 14 | 14 | | SULFATION (BY SIMILARITY). |
| FT | MOD_RES | 15 | 15 | | SULFATION (BY SIMILARITY). |
| SQL | SEQUENCE | 352 AA; | 40527 MW; | F4E2P47135AF58A CRC64; | |

Alignment Scores:

| | | | |
|------------------------|----------|---------------|-----|
| Pred. No.: | 5,57e-14 | Length: | 352 |
| Score: | 152.00 | Matches: | 31 |
| Percent Similarity: | 100.00% | Conservative: | 0 |
| Best Local Similarity: | 100.00% | Mismatches: | 0 |
| Query Match: | 92.68% | Indels: | 0 |
| DB: | 1 | Gaps: | 0 |

US-09-938-703-3_COPY_792_885 (1-94) x CKR5_PONPY (1-352)

| | | | |
|----|-----|--|-----|
| QY | 2 | TAAAGATAGTCATCTTGCGGCTGTCTCCGCTGCTGCATGGTCACTGCTACTCG | 61 |
| DB | 196 | LeuLyslleValIleLeuGlyLeuValLeuProLeuLeuValMetValIleCysTyrSer | 215 |
| QY | 62 | GGATCCCTAAAACTCTGCTCGGTGTCGAAT | 94 |
| DB | 216 | GlylleLeuLysThrLeuLeuArgCysArgAsn | 226 |

RESULT 9
CKR5_PYGBI

| | | | | |
|----|--|-----------|------|---------|
| ID | CKR5_PYGBI | STANDARD; | PRT; | 352 AA. |
| DC | O97880; | | | |
| DT | 30-MAY-2000 (Rel. 39, Created) | | | |
| DT | 30-MAY-2000 (Rel. 39, Last sequence update) | | | |
| DT | 15-JUN-2002 (Rel. 41, Last annotation update) | | | |
| DN | C-C chemokine receptor type 5 (C-C CKR-5) (CCR-5) (CCR5). | | | |
| GN | OS Pygathrix bieti (Black snub-nosed monkey) (Rhinopthecus bieti); | | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutelestomi; | | | |
| OC | Mammalia; Buthera; Primates; Catarrhini; Cercopithecidae; Colobinae; | | | |
| OX | Pygathrix. | | | |
| OC | NCBI_TaxID=61621; | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| EX | MEDLINE=99416438; PubMed=10486970; | | | |
| RA | Zhang Y.-W., Ryder O.A., Zhang I.-P.; | | | |
| FA | "sequence evolution of the CCR5 chemokine receptor gene in primates."; | | | |
| RL | Mol. Biol. Evol. 16:1145-1154(1999). | | | |
| CC | -! FUNCTION: RECEPTOR FOR A C-C TYPE CHEMOKINE. BINDS TO MIP-1-ALPHA, | | | |
| CC | MIP-1-BETA AND RANTES AND SUBSEQUENTLY TRANSDUCES A SIGNAL BY | | | |
| CC | INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL. MAY PLAY A ROLE | | | |
| CC | IN THE CONTROL OF GRANULOCYTIC LINEAGE PROLIFERATION OR | | | |
| CC | DIFFERENTIATION. | | | |
| CC | -! SUBCELLULAR LOCATION: Integral membrane protein. | | | |
| CC | -! SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS. | | | |

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial purposes without prior written permission from the European Bioinformatics Institute is prohibited.

Db 196 LeuysileValleLeuGlyLeuValleProLeuLeuMetValleCysTyrSer 215
 Qy 62 GGAATCTAAACACTCTGCTGCGTGCAGAAAT 94
 Db 216 GlylleLeuLysThrLeuLeuArgCysArgAsn 226
 RESULT 13
 CKR5_MACMU STANDARD; PRT; 352 AA.
 AC P79436; O02746;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE C-C chemokine receptor type 5 (C-C CKR-5) (CCR-5) (CCR5).
 GN CCR5 OR CMKBR5.
 OS Macaca mulatta (Rhesus macaque),
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey), and
 OS Macaca nemestrina (Pig-tailed macaque).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecinae; Macaca.
 OC NCBI_TaxID=9544; 9541, 9545;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX SPECIES=M.mulatta;
 RX MEDLINE=97184592; PubMed=9032394;
 RA Marcon L., Choe H., Martin K.A., Farzan M., Ponath P.D., Wu L.,
 RA Newman W., Gerard N., Gerard C., Sodroski J.;
 RT "Utilization of C-C chemokine receptor 5 by the envelope
 RT glycoproteins of a pathogenic simian immunodeficiency virus,
 RT SIYmac239.";
 RL J. Virol. 71:2522-2527(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX SPECIES=M.mulatta; STRAIN=Indian macaque;
 RX MEDLINE=97213934; PubMed=9060623;
 RA Chen Z., Zhou P., Ho D.D., Landau N.R., Marx P.A.;
 RA "Genetically divergent strains of simian immunodeficiency virus use
 RT CCR5 as a coreceptor for entry.";
 RL J. Virol. 71:2705-2714(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX SPECIES=M.mulatta;
 RX MEDLINE=21354176; PubMed=11461684;
 RA Margulies B.J., Haer D.A., Clements J.E.;
 RT "Identification and comparison of eleven rhesus macaque chemokine
 RT receptors.";
 RL AIDS Res. Hum. Retroviruses 17:981-986(2001).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX SPECIES=M.mulatta, M.fascicularis, and M.nemestrina;
 RX MEDLINE=97268687; PubMed=9108095;
 RA Edinger A.L., Amodee A., Miller K., Doranz B.J., Endres M.,
 RA Sharon M., Sansom M., Lu Z.H., Clements J.E., Murphy-Corb M.,
 RA Peiper S.C., Parmentier M., Broder C.C., Doms R.W.;
 RT "Differential utilization of CCR5 by macrophage and T cell tropic
 RT simian immunodeficiency virus strains.";
 RL Proc. Natl. Acad. Sci. U.S.A. 94:4005-4010(1997).
 CC -1- FUNCTION: RECEPTOR FOR A C-C TYPE CHEMOKINE. BINDS TO MIP-1-ALPHA,
 CC MIP-1-BETA AND RANTES AND SUBSEQUENTLY TRANSDUCES A SIGNAL BY
 CC INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL. MAY PLAY A ROLE
 CC IN THE CONTROL OF GRANULOCYTIC LINEAGE PROLIFERATION OR
 CC DIFFERENTIATION.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

or send an email to license@isb-sib.ch.
 CC
 CC EMBL; U77672; AAC51109.1; -;
 DR EMBL; U73739; AAC51158.1; -;
 DR EMBL; U96762; AAC34132.1; -;
 DR EMBL; AF005660; AAB62554.1; -;
 DR EMBL; AF005661; AAB62555.1; -;
 DR EMBL; AF005662; AAB62556.1; -;
 DR InterPro; IPR000276; GPCR_Rhodpsn.
 DR Pfam; PF00001; 7tm_1; 1.
 DR PRINTS; PR00237; GPCRHHODPSN.
 DR PROSITE; PS00237; G-PROTEIN_RECEP_F1_1; 1.
 DR PROSITE; PS0262; G-PROTEIN_RECEP_F1_2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation.
 FT DOMAIN 1 30 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 31 58 1 (POTENTIAL).
 FT DOMAIN 59 68 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 69 89 2 (POTENTIAL).
 FT DOMAIN 90 102 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 103 124 3 (POTENTIAL).
 FT DOMAIN 125 141 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 142 166 4 (POTENTIAL).
 FT DOMAIN 167 198 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 199 218 5 (POTENTIAL).
 FT DOMAIN 219 235 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 236 260 6 (POTENTIAL).
 FT DOMAIN 261 277 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 278 301 7 (POTENTIAL).
 FT DOMAIN 302 352 CYTOPLASMIC (POTENTIAL).
 FT DISULFID 101 178 BY SIMILARITY.
 FT MOD_RES 3 3 SULFATION (BY SIMILARITY).
 FT MOD_RES 10 10 SULFATION (BY SIMILARITY).
 FT MOD_RES 14 14 SULFATION (BY SIMILARITY).
 FT MOD_RES 15 15 SULFATION (BY SIMILARITY).
 FT CARBOHYD 268 268 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CONFLICT 241 241 M -> I (IN REF. 3).
 FT CONFLICT 292 292 I -> M (IN REF. 3).
 SQ SEQUENCE 352 AA; 40507 MW; 58B96C85909FACE2 CRC64;
 Alignment Scores:
 Pred. No.: 1.51e-13 Length: 352
 Score: 149.00 Matches: 30
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 96.77% Mismatches: 0
 Query Match: 90.85% Indels: 0
 DB: 1 Gaps: 0
 US-09-938-703-3_COPY_792_885 (1-94) x CKR5_MACMU (1-352)
 Qy 2 TTAAGATAGTCATCTGGGGCTGCTGCGCTGCTGCTCATGTCATGTCGTCG 61
 Db 196 LeuLysMetValleLeuGlyLeuValleProLeuLeuValleMetValleCysTyrSer 215
 Qy 62 GGAATCTAAACACTCTGCTGCGTGCAGAAAT 94
 Db 216 GlylleLeuLysThrLeuLeuArgCysArgAsn 226
 RESULT 14
 CKR2_MACMU STANDARD; PRT; 360 AA.
 AC Q18793;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE C-C chemokine receptor type 2 (C-C CKR-2) (CCR-2) (CCR2)
 DE (Monocyte chemoattractant protein 1 receptor) (MCP-1-R) (CCR2).
 GN CCR2 OR CMKBR2.
 OS Macaca mulatta (Rhesus macaque).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecinae; Macaca.
 OC NCBI_TaxID=9544;
 RN [1]

Search completed: June 3, 2003, 18:46:03
Job time : 3.52995 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 3, 2003, 18:42:09 ; Search time 3.30469 Seconds
(without alignments)
5468.980 Million cell updates/sec

Title: US-09-938-703-3_COPY_792_885

Perfect score: 164
Sequence: 1 ATTAAGATAGTATCTTGG.....CTCTGCTTCGGTGCAGAAAT 94

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delep 6.0 , Delext 7.0

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 566448

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_n2p_model -DEV=xlp
-Q/cgr2_1/USPTO_spool/US9938703/runat_03062003_161407_22855/app_query.fasta_1.1870
-DB=pir73 -QFWT=Eastan -SUFFIX=rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFWT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US9938703 -RCGN_1_1_89_ernut_03062003_161407_22855 -NCPU=3
-NO_MAP -LARGESCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPOP=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: PIR73:

1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Query Match | Length | ID | Description |
|------------|-------------|--------|--------------|--------------------|
| 1 | 152 | 92.7 | 352 2 A43113 | chemokine (C-C) re |
| 2 | 134 | 81.7 | 360 2 JC2443 | chemokine (C-C) re |
| 3 | 134 | 81.7 | 374 2 I38450 | chemokine (C-C) re |
| 4 | 122 | 74.4 | 359 2 I49341 | MIP-1 alpha recept |
| 5 | 109 | 66.5 | 355 2 G02436 | chemokine (C-C) re |
| 6 | 106 | 64.6 | 355 2 A45177 | chemokine (C-C) re |
| 7 | 105 | 64.0 | 355 2 I49339 | macrophage inflam |
| 8 | 95 | 57.9 | 350 2 A57160 | chemokine (C-C) re |
| 9 | 88 | 53.7 | 354 2 I58186 | probable G protein |
| 10 | 84 | 51.2 | 383 2 S55594 | G protein-coupled |
| 11 | 81 | 49.4 | 356 2 I49340 | MIP-1 alpha recept |
| 12 | 80 | 48.8 | 360 2 JC4587 | chemokine (C-C) re |
| 13 | 79 | 48.2 | 355 2 JC4304 | orphan G protein-c |
| 14 | 76 | 46.3 | 378 2 A45680 | G protein-coupled |

| | | | | | |
|----|----|------|-----|----------|--------------------|
| 15 | 76 | 46.3 | 378 | 2 B55735 | lymphocyte-specifi |
| 16 | 75 | 45.7 | 355 | 2 JC5067 | G protein-coupled |
| 17 | 74 | 45.1 | 369 | 2 JC5068 | G protein-coupled |
| 18 | 71 | 43.3 | 378 | 2 A55735 | G protein-coupled |
| 19 | 68 | 41.5 | 367 | 2 JB0349 | interferon-inducib |
| 20 | 66 | 40.2 | 350 | 2 JN0621 | G protein-coupled |
| 21 | 66 | 40.2 | 354 | 2 B55733 | G protein-coupled |
| 22 | 66 | 40.2 | 354 | 2 A36689 | interleukin-8 rece |
| 23 | 66 | 40.2 | 355 | 2 JQ1231 | interleukin-8 rece |
| 24 | 66 | 40.2 | 363 | 2 I48261 | angiotensin II rec |
| 25 | 66 | 40.2 | 363 | 2 A49092 | angiotensin II rec |
| 26 | 66 | 40.2 | 363 | 2 JC2543 | angiotensin II rec |
| 27 | 64 | 39.0 | 356 | 2 S42096 | interleukin-8 rece |
| 28 | 64 | 39.0 | 358 | 2 A3752 | interleukin-8 rece |
| 29 | 63 | 38.4 | 359 | 2 A48921 | interleukin-8 rece |
| 30 | 63 | 38.4 | 360 | 2 A53611 | interleukin-8 rece |
| 31 | 63 | 38.4 | 361 | 2 B45680 | G protein-coupled |
| 32 | 62 | 37.8 | 374 | 2 S32785 | G protein-coupled |
| 33 | 60 | 36.6 | 350 | 2 A39445 | interleukin-8 rece |
| 34 | 60 | 36.6 | 375 | 2 JC5509 | G protein-coupled |
| 35 | 59 | 36.0 | 344 | 2 T09508 | intron 17 purinerg |
| 36 | 59 | 36.0 | 358 | 2 G02670 | IL8-related recept |
| 37 | 59 | 36.0 | 374 | 2 S42628 | G protein-coupled |
| 38 | 59 | 36.0 | 375 | 2 JC5069 | G protein-coupled |
| 39 | 57 | 34.8 | 314 | 2 T46279 | hypothetical prote |
| 40 | 57 | 34.8 | 366 | 2 I48469 | Mel-1a melatonin r |
| 41 | 57 | 34.8 | 370 | 2 JC5549 | heptahelical P215- |
| 42 | 57 | 34.8 | 482 | 2 S65766 | G protein-coupled |
| 43 | 56 | 34.1 | 240 | 2 S55480 | hypothetical prote |
| 44 | 56 | 34.1 | 350 | 1 A37963 | complement C5a ana |
| 45 | 56 | 34.1 | 380 | 2 JC2434 | kappa opioid recep |

ALIGNMENTS

RESULT 1

A43113

chemokine (C-C) receptor 5 - human

N:Alternate names: C-CR-5, CCR5

C:Species: Homo sapiens (man)

C>Date: 12-Jul-1996 #sequence,revision 12-Jul-1996 #text_change 20-Jun-2000

C:Accession: A43113; S71808; A58834; A58832; G02653; A58833

R:Samson, M.; Labbe, O.; Mollereau, C.; Vassart, G.; Parmentier, M.

Biochemistry 35, 3362-3367, 1996

A>Title: Molecular cloning and functional expression of a new human CC-chemokine rece

A:Reference number: A43113; MUID:96241590; PMID:8639485

A:Accession: A43113

A:Molecule type: mRNA

A:Residues: 1-352 <SAML>

A:Cross-references: GB:X01492; NID:gl262810; PIDN:CAA62796.1; PID:gl262811

R:Samson, M.; Libert, F.; Doranz, B.J.; Rucker, J.; Lienard, C.; Farber, C.M.; Sarag

M.; Imal, T.; Rana, S.; Yi, Y.; Smyth, R.J.; Collman, R.G.; Doms, R.W.; Vassart, G.;

Nature 382, 722-725, 1996

A>Title: Resistance to HIV-1 infection in caucasian individuals bearing mutant allele

A:Reference number: S71808; MUID:96345670; PMID:8751444

A:Accession: S71808

A>Status: nucleic acid sequence not shown; not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 182-206;207-230 <SAM2>

A:Accession: A58834

A>Status: nucleic acid sequence not shown; not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-184, 'IKDSHLGAGPAAACHGILLGNPKNSAVSK' <SAM3>

A:Cross-references: GB:X99393; NID:gl524062; PIDN:CA467767.1; PID:gl524063

A>Note: This frameshift mutation results in a non-functional receptor but confers a c

nd may have had a selective advantage by conferring resistance to Versinia plague inf

J. Combadiere, C.; Ahuja, S.K.; Tiffany, H.L.; Murphy, P.M.

T. Leukoc. Biol. 60, 147-152, 1996

A>Title: Cloning and functional expression of CC CR5, a human monocyte CC chemokine

A:Reference number: A58832; MUID:96295970; PMID:8699119

A:Accession: A58832

A:Molecule type: mRNA

A:Residues: 1-352 <COML>

A;Cross-references: DDBJ:D29984; NID:9531246; PIDN:BA06253.1; PID:9531247
F;Charo, I.F.; Myers, S.J.; Herman, A.; Franci, C.; Connolly, A.J.; Coughlin, S.R.
Proc. Natl. Acad. Sci. U.S.A. 91, 2752-2756, 1994
A;Title: Molecular cloning and functional expression of two monocyte chemoattractant
A;Reference number: A53477; MUID:94195821; PMID:8146186
A;Accession: I38463
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-360 <RES>
A;Cross-references: EMBL:U03905; NID:9472557; PIDN:AAA19120.1; PID:9472558
C;Genetics:
A;Gene: GDB:CMKBR2
A;Cross-references: GDB:337364; OMIM:601267
A;Map position: 3p21-3p21
C;Superfamily: vertebrate rhodopsin
C;Keywords: alternative splicing; G protein-coupled receptor; glycoprotein; transmembrane
F;43-70/Domain: transmembrane #status predicted <TM>
F;115-136/Domain: transmembrane #status predicted <TM2>
F;154-178/Domain: transmembrane #status predicted <TM4>
F;207-226/Domain: transmembrane #status predicted <TM>
F;244-268/Domain: transmembrane #status predicted <TM6>
F;287-309/Domain: transmembrane #status predicted <TM7>
F;4/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;113-190/Disulfide bonds: #status predicted

Alignment Scores:

| Pred. No.: | 5.5e-11 | Length: | 360 |
|------------------------|---------|---------------|-----|
| Score: | 134.00 | Matches: | 26 |
| Percent Similarity: | 100.00% | Conservative: | 1 |
| Best Local Similarity: | 96.30% | Mismatches: | 0 |
| Query Match: | 81.71% | Indels: | 0 |
| DB: | 2 | Gaps: | 0 |

US-09-938-703-3_COPY_792_885 (1-94) x JC2443 (1-360)

Qy 14 ATCTGGGCGTGGTCTGGCGTGTTCATGCTACTCTCGGAATCCTAAAA 73
|||||
Db 208 ILEucluyLeuValleuProleuLeuIleMetValleuCystyrSerGlyIleLeuLys 227
|||||
Qy 74 ACTCTGCTCTGGTGTGGAAT 94
|||||
Db 228 ThrLeuLeuArgCysArgAsn 234
|||||

RESULT 3
I38450
chemokine (C-C) receptor 2, splice form A - human
N;Alternate names: C-C CKR-2; monocyte chemoattractant protein 1 receptor; monocyte
C;Species: Homo sapiens (man)
C;Date: 16-Feb-1995 #sequence_revision 16-Feb-1996 #text_change 13-Aug-1999
C;Accession: I38450
F;Charo, I.F.; Myers, S.J.; Herman, A.; Franci, C.; Connolly, A.J.; Coughlin, S.R.
Proc. Natl. Acad. Sci. U.S.A. 91, 2752-2756, 1994
A;Title: Molecular cloning and functional expression of two monocyte chemoattractant
A;Reference number: A53477; MUID:94195821; PMID:8146186
A;Accession: I38450
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-374 <RES>
A;Cross-references: EMBL:U03882; NID:9472555; PIDN:AAA19119.1; PID:9472556
C;Genetics:
A;Gene: GDB:CMKBR2
A;Cross-references: GDB:337364; OMIM:601267
A;Map position: 3p21-3p21
C;Superfamily: vertebrate rhodopsin
C;Keywords: alternative splicing; G protein-coupled receptor; glycoprotein; transmembrane
F;44-66/Domain: transmembrane #status predicted <TM1>
F;79-99/Domain: transmembrane #status predicted <TM2>
F;115-136/Domain: transmembrane #status predicted <TM3>
F;154-178/Domain: transmembrane #status predicted <TM4>
F;208-226/Domain: transmembrane #status predicted <TM5>
F;244-265/Domain: transmembrane #status predicted <TM6>
F;292-309/Domain: transmembrane #status predicted <TM7>


```

Alignment Scores:      5.9e-07      355
Pred. No.:           106.00      21
Score:               83.87%
Percent Similarity:   67.74%
Best Local Similarity: 5
Mismatch:             0
Query Match:         64.63%
DB:                   2
Gaps:                0

US-09-938-703-3_COPY_792_885 (1-94) x A45177 (1-355)

QY    2 TTAAGACATAGTCATCTGGGCTGGTCCCGCGCTGTGTTGCATGTCATCTGCTACTCG 61
      |||||::: :: | |||||:::|||||:::|||||:::|||||:::|||||:::
Db    201 LeuLysLeuAsnLeuPheGlyLeuValLeuProLeuValMetIleCysrYrHr 220

QY    62 GGAAATCCATAAAGACTCTGCTCGGTGTCGAAT 94
      |||||:::|||| | |||||:::||||
Db    221 GlyIleIeuLysIleLeuLeuArgArgProAsn 231

RESULT 7
I49339
macrophage inflammatory protein-1 alpha receptor - mouse
C:Species: Mus musculus (house mouse)
C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 13-Aug-1999
C:Accession: I49339
R:Gao, J.L.; Murphy, P.M.
J. Biol. Chem. 270, 17494-17501, 1995
A:Title: Cloning and differential tissue-specific expression of three mouse beta chemokine genes
A:Reference number: I49339; MUID:95340546; PMID:7542241
A:Accession: I49339
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-355 <RES>
A:Cross-references: EMBL:U28404; NID:g881547; PIDN:AAA89153.1; PID:g881548
C:Superfamily: vertebrate rhodopsin

```

158186
 probable G protein-coupled receptor - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 21-Jul-2000
 C:Accession: 158186
 R:Harrison, J.K.; Barber, C.M.; Lynch, K.R.
 Neurosci. Lett. 169, 85-89, 1994
 A:Title: cDNA cloning of a G-protein-coupled receptor expressed in rat spinal cord and
 A:Reference number: 158186; MUID:94323113; PMID:8047298
 A:Accession: 158186
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-354 <RES>
 A:Cross-references: EMBL:U04808; NID:g2558635; PIDN:AA887093.1; PID:g439861
 C:Superfamily: vertebrate rhodopsin
 C:Keywords: G protein-coupled receptor

| | | | |
|------------------------|---------|---------------|-----|
| Alignment Scores: | 0.00023 | Length: | 354 |
| Pred. No.: | 88.00 | Matches: | 16 |
| Score: | 73.33% | Conservative: | 6 |
| Percent Similarity: | 53.33% | Mismatches: | 8 |
| Best Local Similarity: | 53.66% | Indels: | 0 |
| Query Match: | 2 | Gaps: | 0 |
| DB: | | | |

US-09-938-703-3 COPY 792 885 (1-94) x I49340 (1-356)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2003, 19:19:16 ; Search time 44.5732 Seconds
(without alignments)
993.874 Million cell updates/sec

Title: US-09-938-703-6

Perfect score: 1122
Sequence: 1 MDYQVSPYIDINYYTSEPC.....AACGHLLGNPKNSASYSK 215

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL_21.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_virus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | ID | Description |
|------------|-------|---------------|--------|----|--------------------|
| 1 | 1122 | 100.0 | 215 | 4 | 075303 homo sapien |
| 2 | 947 | 84.4 | 352 | 6 | Q95NC5 hylobates s |
| 3 | 945 | 84.2 | 352 | 6 | 018772 pan troglod |
| 4 | 940 | 83.8 | 352 | 6 | 018771 pan troglod |
| 5 | 940 | 83.8 | 352 | 6 | Q95NC0 hylobates m |
| 6 | 939 | 83.7 | 352 | 6 | Q9XS99 gorilla gor |
| 7 | 938 | 83.6 | 352 | 6 | Q9IV50 pan troglod |
| 8 | 938 | 83.6 | 352 | 6 | Q95NC8 colobus pol |
| 9 | 938 | 83.6 | 352 | 6 | Q95NC6 trachypithe |
| 10 | 935 | 83.3 | 352 | 6 | Q9MZA3 hylobates a |
| 11 | 935 | 83.3 | 352 | 6 | 097962 pygathrix a |
| 12 | 934 | 83.2 | 352 | 6 | Q9XT14 colobus gue |
| 13 | 934 | 83.2 | 352 | 6 | Q95NC7 nasalis lar |
| 14 | 934 | 83.2 | 352 | 6 | Q95NC3 miopithecus |
| 15 | 933 | 83.2 | 352 | 6 | 018770 pan troglod |
| 16 | 933 | 83.2 | 352 | 6 | Q9TV49 cercocobus |

| | | | | | | |
|----|-----|------|-----|---|--------|--------------------|
| 17 | 933 | 83.2 | 352 | 6 | 097975 | 097975 macaca arct |
| 18 | 930 | 82.9 | 352 | 6 | Q9TQX0 | Q9TQX0 cercopithec |
| 19 | 929 | 82.8 | 352 | 6 | Q9XT12 | Q9XT12 cercopithec |
| 20 | 929 | 82.8 | 352 | 6 | Q95NC1 | Q95nc1 theropithec |
| 21 | 927 | 82.6 | 352 | 6 | Q9TV46 | Q9TV46 cercopithec |
| 22 | 927 | 82.6 | 352 | 6 | 077776 | 077776 cercocobus |
| 23 | 927 | 82.6 | 352 | 6 | Q95ND2 | Q95nd2 mandrillus |
| 24 | 926 | 82.5 | 352 | 6 | Q9XT13 | Q9XT13 papio anubi |
| 25 | 925 | 82.4 | 352 | 6 | Q9TSK1 | Q9TSK1 cercopithec |
| 26 | 925 | 82.4 | 352 | 6 | Q9MZA2 | Q9mza2 cercopithec |
| 27 | 925 | 82.4 | 352 | 6 | Q95NE8 | Q95ne8 cercocobus |
| 28 | 925 | 82.4 | 352 | 6 | Q95NE1 | Q95ne1 homo sapien |
| 29 | 923 | 82.3 | 339 | 4 | Q9UN28 | Q9un28 homo sapien |
| 30 | 923 | 82.3 | 339 | 4 | Q9UN23 | Q9un23 homo sapien |
| 31 | 923 | 82.3 | 339 | 4 | Q9UBT9 | Q9ubt9 homo sapien |
| 32 | 923 | 82.3 | 339 | 4 | Q9UBJ7 | Q9ubj7 homo sapien |
| 33 | 922 | 82.2 | 352 | 6 | Q9TV93 | Q9TV93 macaca arct |
| 34 | 922 | 82.2 | 352 | 6 | Q95ND1 | Q95nd1 mandrillus |
| 35 | 922 | 82.2 | 352 | 6 | Q95ND0 | Q95nd0 erythrocebu |
| 36 | 920 | 82.0 | 339 | 4 | Q9UN27 | Q9un27 homo sapien |
| 37 | 920 | 82.0 | 352 | 6 | Q9TV42 | Q9TV42 cercopithec |
| 38 | 919 | 81.9 | 352 | 6 | Q9TV43 | Q9TV43 cercopithec |
| 39 | 919 | 81.9 | 352 | 6 | Q9XS35 | Q9xs35 macaca neme |
| 40 | 917 | 81.7 | 339 | 6 | Q9TQW7 | Q9TQW7 pan troglod |
| 41 | 917 | 81.7 | 339 | 6 | Q9TQW6 | Q9TQW6 pan troglod |
| 42 | 917 | 81.7 | 339 | 6 | Q9TQW4 | Q9TQW4 pan troglod |
| 43 | 916 | 81.6 | 352 | 6 | Q9TV44 | Q9TV44 cercopithec |
| 44 | 916 | 81.6 | 352 | 6 | Q9XT76 | Q9XT76 cercopithec |
| 45 | 915 | 81.6 | 352 | 6 | Q9TV45 | Q9TV45 cercopithec |

ALIGNMENTS

RESULT 1

O75303 PRELIMINARY; PRT; 215 AA.

AC O75303;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE CC-chemokine receptor.
GN CCR-5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Tse L., Ehrenberg P.K., Chang G., Michael N.L.;
RT "Genomic Organization and Functional Characterization of the Complete
RT Transcription Unit for the Chemokine Receptor CCR-5, a Major Entry Co-
RT Receptor for HIV-1.";
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF009962; AAC23944.1; -
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1.1.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
SQ SEQUENCE 215 AA; 23946 MW; 3C9146C76BA16F7 CRC64;

Query Match 100.0%; Score 1122; DB 4; Length 215;
Best Local Similarity 100.0%; Pred. No. 5.8e-97;
Matches 215; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDYQVSPYIDINYYTSEPCQKINVKQIAARLLPPLSLVIFGFGVGNMVLILLINCKR 60
|||||
Db 1 MDYQVSPYIDINYYTSEPCQKINVKQIAARLLPPLSLVIFGFGVGNMVLILLINCKR 60
|||||

Qy 61 LKSMTDIYLINLAISDLFFLLITVPFWAHYAAQWDFNTMCQLLGLYFIQFSGIFFI 120
|||||
Db 61 LKSMTDIYLINLAISDLFFLLITVPFWAHYAAQWDFNTMCQLLGLYFIQFSGIFFI 120
|||||

QY 121 LITDRLAVVHAFKARTVFGVTSVITWVAVFASLPGLIFTRSQEGLHYTCSS 180
 DB 121 LITDRLAVVHAFKARTVFGVTSVITWVAVFASLPGLIFTRSQEGLHYTCSS 180
 QY 181 HFPYKDSHGAPAAACHGHLGNPKNSASYSK 215
 DB 181 HFPYKDSHGAPAAACHGHLGNPKNSASYSK 215

RESULT 2

Q95NC5 PRELIMINARY; PRT; 352 AA.
 ID Q95NC5
 AC Q95NC5;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE C-C chemokine receptor 5.
 GN CCR5.
 OS Hylobates syndactylus (Siamang) (Symphalangus syndactylus).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Primates; Catarrhini; Hylobatidae; Hylobates.
 OX NCBI_TaxID=9590;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Zhang Y., Ryder O.A., Zhang Y.;
 RT "Sequence comparison of the CCR5 gene in primates and primate
 RT phylogeny."
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF177884; AAK43367.1; -
 DR InterPro: IPR000276; GPCR_Rhodpsn.
 DR Pfam: PF00001; 7tm_1; 1.
 DR PROSITE; PS00237; G_PROTEIN_RECEPT_F1_1; UNKNOWN_1.
 DR PROSITE; PS50262; G_PROTEIN_RECEPT_F1_2; 1.
 KW Receptor.
 SQ SEQUENCE 352 AA; 40508 MW; F4F64B3AD5AF658A CRC64;

Query Match 84.4%; Score 947; DB 6; Length 352;
 Best Local Similarity 98.4%; Pred. No. 1.9e-80;
 Matches 181; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MDYQVSSPIYDINTYSEPCQKINVKQIAARLLPPLYSLVIFGFGVGNMLVILLINCKR 60
 DB 1 MDYQVSSPIYDINTYSEPCQKINVKQIAARLLPPLYSLVIFGFGVGNMLVILLINCKR 60
 QY 61 LKSMTDIYLLNLAIISDLFFLLTPVFWAHYAAQWDFGNTMCLLGLYFGFSGIFII 120
 DB 61 LKSMTDIYLLNLAIISDLFFLLTPVFWAHYAAQWDFGNTMCLLGLYFGFSGIFII 120
 QY 121 LLTIDRYLAVHAFKARTVFGVTSVITWVAVFASLPGLIFTRSQEGLHYTCSS 180
 DB 121 LLTIDRYLAVHAFKARTVFGVTSVITWVAVFASLPGLIFTRSQEGLHYTCSS 180
 QY 181 HFPY 184
 DB 181 HFPY 184

RESULT 3

ID 018772 PRELIMINARY; PRT; 352 AA.
 AC 018772;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE CCR5 receptor (fragment).
 GN CCR5.
 OS Pan troglodytes (Chimpanzee).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Pan.
 OX NCBI_TaxID=9598;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN-CHCCR5-142a;
 RA Zhang L., Carruthers C.D., He T., Huang Y., Cao Y., Wang G., Hahn B.,
 RA Ho D.D.;
 RT "HIV-1 subtypes, co-receptor usage, and CCR5 polymorphism.";
 RL AIDS Res. Hum. Retroviruses 0:0-0(1997).
 DR EMBL; AF011541; AAB65741.1; -
 DR InterPro: IPR000276; GPCR_Rhodpsn.
 DR Pfam: PF00001; 7tm_1; 1.
 DR PRINTS; PR00237; GPCRHHODPSN.
 DR PROSITE; PS00237; G_PROTEIN_RECEPT_F1_1; UNKNOWN_1.
 DR PROSITE; PS50262; G_PROTEIN_RECEPT_F1_2; 1.
 KW Receptor.
 FT NON_TER 352 352
 SQ SEQUENCE 352 AA; 40598 MW; 39688FA7004C952F CRC64;

Query Match 84.2%; Score 945; DB 6; Length 352;
 Best Local Similarity 98.4%; Pred. No. 2.9e-80;
 Matches 181; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MDYQVSSPIYDINTYSEPCQKINVKQIAARLLPPLYSLVIFGFGVGNMLVILLINCKR 60
 DB 1 MDYQVSSPIYDINTYSEPCQKINVKQIAARLLPPLYSLVIFGFGVGNMLVILLINCKR 60
 QY 61 LKSMTDIYLLNLAIISDLFFLLTPVFWAHYAAQWDFGNTMCLLGLYFGFSGIFII 120
 DB 61 LKSMTDIYLLNLAIISDLFFLLTPVFWAHYAAQWDFGNTMCLLGLYFGFSGIFII 120
 QY 121 LLTIDRYLAVHAFKARTVFGVTSVITWVAVFASLPGLIFTRSQEGLHYTCSS 180
 DB 121 LLTIDRYLAVHAFKARTVFGVTSVITWVAVFASLPGLIFTRSQEGLHYTCSS 180
 QY 181 HFPY 184
 DB 181 HFPY 184

RESULT 4

ID 018771 PRELIMINARY; PRT; 352 AA.
 AC 018771;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE CCR5 receptor (fragment).
 GN CCR5.
 OS Pan troglodytes (Chimpanzee).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Pan.
 OX NCBI_TaxID=9598;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CHCCR5-141a;
 RA Zhang L., Carruthers C.D., He T., Huang Y., Cao Y., Wang G., Hahn B.,
 RA Ho D.D.;
 RT "HIV-1 subtypes, co-receptor usage, and CCR5 polymorphism.";
 RL AIDS Res. Hum. Retroviruses 0:0-0(1997).
 DR EMBL; AF011539; AAB65739.1; -
 DR InterPro: IPR000276; GPCR_Rhodpsn.
 DR Pfam: PF00001; 7tm_1; 1.
 DR PRINTS; PR00237; GPCRHHODPSN.
 DR PROSITE; PS00237; G_PROTEIN_RECEPT_F1_1; UNKNOWN_1.
 DR PROSITE; PS50262; G_PROTEIN_RECEPT_F1_2; 1.
 KW Receptor.
 FT NON_TER 352 352
 SQ SEQUENCE 352 AA; 40466 MW; 3EFFAC7ABAED1D4FB CRC64;

Query Match 83.8%; Score 940; DB 6; Length 352;
 Best Local Similarity 97.8%; Pred. No. 8.6e-80;
 Matches 180; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MDYQVSSPIYDINTYSEPCQKINVKQIAARLLPPLYSLVIFGFGVGNMLVILLINCKR 60
 DB 1 MDYQVSSPIYDINTYSEPCQKINVKQIAARLLPPLYSLVIFGFGVGNMLVILLINCKR 60

QY 61 LKSMDDIYLLNLAIISDLFFLLVFPWAHYAAQWDFGNTMCQLLGLYFIFGFSGIFII 120
 DB 1 LKSMDDIYLLNLAIISDLFFLLVFPWAHYAAQWDFGNTMCQLLGLYFIFGFSGIFII 120
 QY 121 LKSMDDIYLLNLAIISDLFFLLVFPWAHYAAQWDFGNTMCQLLGLYFIFGFSGIFII 180
 DB 121 LKSMDDIYLLNLAIISDLFFLLVFPWAHYAAQWDFGNTMCQLLGLYFIFGFSGIFII 180
 QY 121 LKSMDDIYLLNLAIISDLFFLLVFPWAHYAAQWDFGNTMCQLLGLYFIFGFSGIFII 180
 DB 121 LKSMDDIYLLNLAIISDLFFLLVFPWAHYAAQWDFGNTMCQLLGLYFIFGFSGIFII 180
 QY 181 HFPY 184
 DB 181 HFPY 184

RESULT 5

Q95NC0 PRELIMINARY; PRT; 352 AA.
 AC Q95NC0;
 DT 01-DEC-2001 (TEMBLrel. 19, Created)
 DT 01-DEC-2001 (TEMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TEMBLrel. 20, Last annotation update)
 DE C-C chemokine receptor 5.
 GN CCR5.
 OS Hylobates moloch (silvery gibbon).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hylobatidae; Hylobates.
 OX NCBI_TaxID=81572;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Zhang Y., Ryder O.A., Zhang Y.;
 RT "Sequence comparison of the CCR5 gene in primates and primate phylogeny";
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF177899; AAK43382.1; -;
 DR InterPro; IPR000276; GPCR_Rhodpsn.
 DR Pfam; PF00001; 7tm.1; 1.
 DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
 DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
 KW Receptor.
 SQ SEQUENCE 352 AA; 40436 MW; 9623CA98340CF274 CRC64;

Query Match 83.8%; Score 940; DB 6; Length 352;
 Best Local Similarity 97.8%; Pred. No. 8.6e-80;
 Matches 180; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 1 MDYQVSSPYDINYYTSEPCQKINVKQIAARLLPPLYSLVIFGFGVGNMLVILLINCKR 60
 DB 1 MDYQVSSPYDINYYTSEPCQKINVKQIAARLLPPLYSLVIFGFGVGNMLVILLINCKR 60
 QY 61 LKSMDDIYLLNLAIISDLFFLLVFPWAHYAAQWDFGNTMCQLLGLYFIFGFSGIFII 120
 DB 61 LKSMDDIYLLNLAIISDLFFLLVFPWAHYAAQWDFGNTMCQLLGLYFIFGFSGIFII 120
 QY 121 LKSMDDIYLLNLAIISDLFFLLVFPWAHYAAQWDFGNTMCQLLGLYFIFGFSGIFII 180
 DB 121 LKSMDDIYLLNLAIISDLFFLLVFPWAHYAAQWDFGNTMCQLLGLYFIFGFSGIFII 180
 QY 181 HFPY 184
 DB 181 HFPY 184

RESULT 6

Q9XS99 PRELIMINARY; PRT; 352 AA.
 AC Q9XS99;
 DT 01-NOV-1999 (TEMBLrel. 12, Created)
 DT 01-NOV-1999 (TEMBLrel. 12, Last sequence update)
 DT 01-JUN-2001 (TEMBLrel. 17, Last annotation update)
 DE C-C chemokine receptor 5.
 GN CCR5.
 OS Gorilla gorilla (gorilla).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
 OX NCBI_TaxID=9593;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-GORILLACCR;
 RX MEDLINE=99210133; PubMed=10195758;
 RA Saksena N.K., Wang B., Novembre F.J., Bolton W., Smit T., Lal R.B.;
 RT "Species-specific changes in the CCR5 gene from African and Asian nonhuman primates";
 RL AIDS Res. Hum. Retroviruses 15:479-483(1999).
 DR EMBL; AF105291; AAD20560.1; -;
 DR InterPro; IPR000276; GPCR_Rhodpsn.
 DR Pfam; PF00001; 7tm.1; 1.
 DR PRINTS; PR00237; GPCR_Rhodpsn.
 DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
 DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
 KW Receptor.
 SQ SEQUENCE 352 AA; 40529 MW; 1B68C68FE2E7AD0 CRC64;

Query Match 83.78%; Score 939; DB 6; Length 352;
 Best Local Similarity 97.3%; Pred. No. 1.1e-79;
 Matches 179; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 QY 1 MDYQVSSPYDINYYTSEPCQKINVKQIAARLLPPLYSLVIFGFGVGNMLVILLINCKR 60
 DB 1 MDYQVSSPYDINYYTSEPCQKINVKQIAARLLPPLYSLVIFGFGVGNMLVILLINCKR 60
 QY 61 LKSMDDIYLLNLAIISDLFFLLVFPWAHYAAQWDFGNTMCQLLGLYFIFGFSGIFII 120
 DB 61 LKSMDDIYLLNLAIISDLFFLLVFPWAHYAAQWDFGNTMCQLLGLYFIFGFSGIFII 120
 QY 121 LKSMDDIYLLNLAIISDLFFLLVFPWAHYAAQWDFGNTMCQLLGLYFIFGFSGIFII 180
 DB 121 LKSMDDIYLLNLAIISDLFFLLVFPWAHYAAQWDFGNTMCQLLGLYFIFGFSGIFII 180
 QY 181 HFPY 184
 DB 181 HFPY 184

RESULT 7

Q9TV50 PRELIMINARY; PRT; 352 AA.
 AC Q9TV50;
 DT 01-MAY-2000 (TEMBLrel. 13, Created)
 DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)
 DT 01-JUN-2001 (TEMBLrel. 17, Last annotation update)
 DE CC chemokine receptor type 5.
 GN CCR5.
 OS Pan troglodytes (Chimpanzee).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
 OX NCBI_TaxID=9598;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1410;
 RX MEDLINE=99335215; PubMed=10408730;
 RA Muller-Trutwin M.C., Corbet S., Hansen J., Georges-Courbot M.C.,
 RA Diop O., Rigoulet J., Barre-Sinoussi F., Fomsgaard A.;
 RT "Mutations in CCR5-coding sequences are not associated with HIV carrier status in African nonhuman primates";
 RL AIDS Res. Hum. Retroviruses 15:931-939(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1410;
 RA Mueller-Trutwin M.C., Corbet S., Hansen J., Diop O.,
 RA Georges-Courbot M.C., Barre-Sinoussi F., Fomsgaard A.;
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF035214; AAD44007.1; -;
 DR InterPro; IPR000276; GPCR_Rhodpsn.
 DR Pfam; PF00001; 7tm.1; 1.
 DR PRINTS; PR00237; GPCR_Rhodpsn.
 DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.

```
DR PROSITE; PS50262; G_PROTEIN_RECEP_FL_2; 1.
KW Receptor.
SQ SEQUENCE 352 AA; 40481 MW; 2578A0F2C07AA65 CRC64;

Query Match      83.6%; Score 938; DB 6; Length 352;
Best Local Similarity 96.2%; Pred. No. 1.3e-79;
Matches 178; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDYQSSPIYDINVTSEPCQKINVKQIAARLLPPLYSLVFIFGVGNMVLILLINCKR 60
DB 1 MDYQSSPTDYDIDYITSEPCQKINVKQIAARLLPPLYSLVFIFGVGNMVLILLINCKR 60
QY 61 LKSMTDIYLLNLAIISDLFFLLTPFWAHYAAQWDFGNMTCQLLTGLYFGFSGIFFI 120
DB 61 LKSMTDIYLLNLAIISDLFFLLTPFWAHYAAQWDFGNMTCQLLTGLYFGFSGIFFI 120
QY 121 LITIDRYLAVHAFALKARTVTFGVVTSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
DB 121 LITIDRYLAVHAFALKARTVTFGVVTSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
QY 181 HPFY 184
DB 181 HPFY 184

RESULT 8
Q95NC8 PRELIMINARY; PRT; 352 AA.
AC Q95NC8;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE C-C chemokine receptor 5.
GN CCR5.
OS Colobus polykomos.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae; Colobinae;
OC Colobus.
OX NCBI_TaxID=9572;
RN [1]
RP SEQUENCE FROM N.A.
RA Zhang Y., Ryder O.A., Zhang Y.;
RT "Sequence comparison of the CCR5 gene in primates and primate
phylogeny.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF177881; AAK43364.1; -.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1; 1.
DR PROSITE; PS00237; G_PROTEIN_RECEP_FL_1; UNKNOWN_1.
DR PROSITE; PS50262; G_PROTEIN_RECEP_FL_2; 1.
KW Receptor.
SQ SEQUENCE 352 AA; 40578 MW; 4366F149C3B4938F CRC64;

Query Match      83.6%; Score 938; DB 6; Length 352;
Best Local Similarity 96.2%; Pred. No. 1.3e-79;
Matches 177; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDYQSSPIYDINVTSEPCQKINVKQIAARLLPPLYSLVFIFGVGNMVLILLINCKR 60
DB 1 MDYQSSPTDYDIDYITSEPCQKINVKQIAARLLPPLYSLVFIFGVGNMVLILLINCKR 60
QY 61 LKSMTDIYLLNLAIISDLFFLLTPFWAHYAAQWDFGNMTCQLLTGLYFGFSGIFFI 120
DB 61 LKSMTDIYLLNLAIISDLFFLLTPFWAHYAAQWDFGNMTCQLLTGLYFGFSGIFFI 120
QY 121 LITIDRYLAVHAFALKARTVTFGVVTSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
DB 121 LITIDRYLAVHAFALKARTVTFGVVTSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
QY 181 HPFY 184
DB 181 HPFY 184
```

```
RESULT 9
Q95NC6 PRELIMINARY; PRT; 352 AA.
AC Q95NC6;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE C-C chemokine receptor 5.
GN CCR5.
OS Trachypithecus johnii (hooded leaf monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae; Colobinae;
OC Trachypithecus.
OX NCBI_TaxID=65063;
RN [1]
RP SEQUENCE FROM N.A.
RA Zhang Y., Ryder O.A., Zhang Y.;
RT "Sequence comparison of the CCR5 gene in primates and primate
phylogeny.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF177883; AAK43366.1; -.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1; 1.
DR PROSITE; PS00237; G_PROTEIN_RECEP_FL_1; UNKNOWN_1.
DR PROSITE; PS50262; G_PROTEIN_RECEP_FL_2; 1.
KW Receptor.
SQ SEQUENCE 352 AA; 40462 MW; 52824E032259F7F CRC64;

Query Match      83.6%; Score 938; DB 6; Length 352;
Best Local Similarity 96.2%; Pred. No. 1.3e-79;
Matches 177; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDYQSSPIYDINVTSEPCQKINVKQIAARLLPPLYSLVFIFGVGNMVLILLINCKR 60
DB 1 MDYQSSPTDYDIDYITSEPCQKINVKQIAARLLPPLYSLVFIFGVGNMVLILLINCKR 60
QY 61 LKSMTDIYLLNLAIISDLFFLLTPFWAHYAAQWDFGNMTCQLLTGLYFGFSGIFFI 120
DB 61 LKSMTDIYLLNLAIISDLFFLLTPFWAHYAAQWDFGNMTCQLLTGLYFGFSGIFFI 120
QY 121 LITIDRYLAVHAFALKARTVTFGVVTSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
DB 121 LITIDRYLAVHAFALKARTVTFGVVTSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
QY 181 HPFY 184
DB 181 HPFY 184

RESULT 10
Q9WZA3 PRELIMINARY; PRT; 352 AA.
AC Q9WZA3;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE CC chemokine receptor 5.
GN CCR5.
OS Hylobates agilis unko.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hylobatidae; Hylobates.
OX NCBI_TaxID=9583;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-20317091; PubMed-10747879;
RA Mummidi S., Bamshad M., Ahuja S.S., Gonzalez E., Feuillet P.M.,
RA Begum K., Galvis M.C., Kosteki V., Valente A.J., Murthy K.K.,
RA Haro L., Dolan M.J., Allan J.S., Ahuja S.K.;
RT "Evolution of human and non-human primate CC chemokine receptor 5 gene
and mRNA. Potential roles for haplotype and mRNA diversity,
RT differential haplotype-specific transcriptional activity, and altered
RT transcription factor binding to polymorphic nucleotides in the
```

RT pathogenesis of HIV-1 and simian immunodeficiency virus.";

RL J. Biol. Chem. 275:18946-18961(2000).
 DR EMBL; AF252551; AAF87981.1; -
 DR InterPro: IPR000276; GPCR_Rhodpsn.
 DR Pfam: PF00001; 7tm.1; 1
 DR PRINTS; PR00237; GPCRHHODPSN
 DR PROSITE; PS00237; G_PROTEIN_RECEP_FL_1; UNKNOWN_1.
 DR PROSITE; PS0262; G_PROTEIN_RECEP_FL_2; 1.
 KW Receptor.
 SQ SEQUENCE 352 AA; 40293 MW; BAF8279ABAB5309 CRC64;

Query Match 83.3%; Score 935; DB 6; Length 352;
 Best Local Similarity 96.7%; Pred. No. 2.5e-79;
 Matches 178; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 MDYQSSPTDYDNYTSEPCQKINVKQIAARLLPPLYSLVIFGVGNMLVILLINCKR 60
 DB 1 MDYQSSPTDYDNYTSEPCQKINVKQIAARLLPPLYSLVIFGVGNMLVILLINCKR 60
 QY 61 LKSMTDIYLLNLAISDLFFLLTPFWAHYAAQWDFGNTMCQLLTGLYFIFGFSGIFII 120
 DB 61 LKSMTDIYLLNLAISDLFFLLTPFWAHYAAQWDFGNTMCQLLTGLYFIFGFSGIFII 120
 QY 121 LTTDRYLAVVHAFKARTVFGVTSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
 DB 121 LTTDRYLAVVHAFKARTVFGVTSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
 QY 181 HPFY 184
 DB 181 HPFY 184

RESULT 11

O97962 ID O97962 PRELIMINARY; PRT; 352 AA.
 AC O97962;
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE C-C chemokine receptor type 5.
 GN CCR5.
 OS Pygathrix avunculus (Tonkin snub-nosed monkey), and
 OS Pygathrix roxellana (golden snub-nosed monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Colobinae;
 OC Pygathrix.
 OC NCBI_TaxID=66062, 61622;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Zhang Y.-W., Zhang Y.-P.;
 RT "Sequence evolution of chemokine receptor CCR5 gene in primates.";
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF075447; AAD19859.1; -
 DR EMBL; AF075444; AAD19856.1; -
 DR InterPro: IPR000276; GPCR_Rhodpsn.
 DR Pfam: PF00001; 7tm.1; 1.
 DR PRINTS; PR00237; GPCRHHODPSN
 DR PROSITE; PS00237; G_PROTEIN_RECEP_FL_1; UNKNOWN_1.
 DR PROSITE; PS0262; G_PROTEIN_RECEP_FL_2; 1.
 KW Receptor.
 SQ SEQUENCE 352 AA; 40482 MW; 037CFA9E12E532F3 CRC64;

Query Match 83.3%; Score 935; DB 6; Length 352;
 Best Local Similarity 95.7%; Pred. No. 2.5e-79;
 Matches 176; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDYQSSPTDYDNYTSEPCQKINVKQIAARLLPPLYSLVIFGVGNMLVILLINCKR 60
 DB 1 MDYQSSPTDYDNYTSEPCQKINVKQIAARLLPPLYSLVIFGVGNMLVILLINCKR 60
 QY 61 LKSMTDIYLLNLAISDLFFLLTPFWAHYAAQWDFGNTMCQLLTGLYFIFGFSGIFII 120
 DB 61 LKSMTDIYLLNLAISDLFFLLTPFWAHYAAQWDFGNTMCQLLTGLYFIFGFSGIFII 120

QY 121 LTTDRYLAVVHAFKARTVFGVTSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
 DB 121 LTTDRYLAVVHAFKARTVFGVTSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
 QY 181 HPFY 184
 DB 181 HPFY 184

RESULT 12

O9XT14 ID O9XT14 PRELIMINARY; PRT; 352 AA.
 AC O9XT14;
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE CC chemokine receptor 5.
 GN CCR5.
 OS Colobus guereza (Black-and-white colobus monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Colobinae;
 OC Colobus.
 OC NCBI_TaxID=33548;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Spearman P.W., Mburu D.N., Graham B.S.;
 RT "Differential utilization of CCR5 molecules from three East African
 RL Simian species by the HIV-1 envelope glycoprotein.";
 RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF141639; AAD32684.1; -
 DR InterPro: IPR000276; GPCR_Rhodpsn.
 DR Pfam: PF00001; 7tm.1; 1.
 DR PRINTS; PR00237; GPCRHHODPSN
 DR PROSITE; PS00237; G_PROTEIN_RECEP_FL_1; UNKNOWN_1.
 DR PROSITE; PS0262; G_PROTEIN_RECEP_FL_2; 1.
 KW Receptor.
 SQ SEQUENCE 352 AA; 40550 MW; 98078EF04D34DB36 CRC64;

Query Match 83.2%; Score 934; DB 6; Length 352;
 Best Local Similarity 95.7%; Pred. No. 3.1e-79;
 Matches 176; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY 1 MDYQSSPTDYDNYTSEPCQKINVKQIAARLLPPLYSLVIFGVGNMLVILLINCKR 60
 DB 1 MDYQSSPTDYDNYTSEPCQKINVKQIAARLLPPLYSLVIFGVGNMLVILLINCKR 60
 QY 61 LKSMTDIYLLNLAISDLFFLLTPFWAHYAAQWDFGNTMCQLLTGLYFIFGFSGIFII 120
 DB 61 LKSMTDIYLLNLAISDLFFLLTPFWAHYAAQWDFGNTMCQLLTGLYFIFGFSGIFII 120
 QY 121 LTTDRYLAVVHAFKARTVFGVTSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
 DB 121 LTTDRYLAVVHAFKARTVFGVTSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
 QY 181 HPFY 184
 DB 181 HPFY 184

RESULT 13

O95NC7 ID O95NC7 PRELIMINARY; PRT; 352 AA.
 AC O95NC7;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE C-C chemokine receptor 5.
 GN CCR5.
 OS Nasalis larvatus (Proboscis monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Colobinae;
 OC Nasalis.

```

OX NCBI_TaxID=43780;
RN [1]
RP SEQUENCE FROM N.A.
RA Zhang Y., Ryder O.A., Zhang Y.;
RT *Sequence comparison of the CCR5 gene in primates and primate
RL phylogeny.*;
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF177882; AAK43365.1; -
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam: PF00001; 7tm_1; 1
DR PROSITE: PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE: PS00262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 352 AA; 40537 MW; 51F6F1486E35938E CRC64;

Query Match 83.2%; Score 934; DB 6; Length 352;
Best Local Similarity 95.7%; Pred. No. 3.1e-79;
Matches 176; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDYQSSPTIDINYTSPEQCKINVKQIAARLLPPLYSLVIFGFGVGNMLVILLINCKR 60
DB 1 MDYQSSPTIDIDYTSPEQCKVNVKQIAARLLPPLYSLVIFGFGVGNMLVILLINCKR 60
QY 61 LKSMTDIYLLNLAIISDLFLLTPFWAHYAAQWDFGNTMCLLTGLYFGFGSIFII 120
DB 61 LKSMTDIYLLNLAIISDLFLLTPFWAHYAAQWDFGNTMCLLTGLYFGFGSIFII 120
QY 121 LTTIDRYLAVHAFALKARTVTFGVTSVITWVAVFASLPFGIIFTRSQEGLHYTCSS 180
DB 121 LTTIDRYLAVHAFALKARTVTFGVTSVITWVAVFASLPFGIIFTRSQEGLHYTCSS 180
QY 181 HFPY 184
DB 181 HFPY 184

RESULT 14
Q95NC3 PRELIMINARY; PRT; 352 AA.
AC Q95NC3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE C-C chemokine receptor 5.
GN CCR5.
OS Miopithecus talapoin (Talapoin) (Cercopithecus talapoin).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciinae; Miopithecus.
OX NCBI_TaxID=36231;
RN [1]
RP SEQUENCE FROM N.A.
RA Zhang Y., Ryder O.A., Zhang Y.;
RT *Sequence comparison of the CCR5 gene in primates and primate
RL phylogeny.*;
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF177886; AAK43369.1; -
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam: PF00001; 7tm_1; 1
DR PROSITE: PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE: PS00262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 352 AA; 40546 MW; 6464152F3E566AE5 CRC64;

Query Match 83.2%; Score 934; DB 6; Length 352;
Best Local Similarity 96.2%; Pred. No. 3.1e-79;
Matches 177; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 MDYQSSPTIDINYTSPEQCKINVKQIAARLLPPLYSLVIFGFGVGNMLVILLINCKR 60
DB 1 MDYQSSPTIDINYTSPEQCKINVKQIAARLLPPLYSLVIFGFGVGNMLVILLINCKR 60
QY 61 LKSMTDIYLLNLAIISDLFLLTPFWAHYAAQWDFGNTMCLLTGLYFGFGSIFII 120
DB 61 LKSMTDIYLLNLAIISDLFLLTPFWAHYAAQWDFGNTMCLLTGLYFGFGSIFII 120
QY 121 LTTIDRYLAVHAFALKARTVTFGVTSVITWVAVFASLPFGIIFTRSQEGLHYTCSS 180
DB 121 LTTIDRYLAVHAFALKARTVTFGVTSVITWVAVFASLPFGIIFTRSQEGLHYTCSS 180
QY 181 HFPY 184
DB 181 HFPY 184

RESULT 15
O18770 PRELIMINARY; PRT; 352 AA.
ID O18770;
AC O18770;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE CCR5 receptor (fragment).
GN CCR5.
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.
OX NCBI_TaxID=9598;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MACCS-140A;
RA Zhang L., Carruthers C.D., He T., Huang Y., Cao Y., Wang G., Hahn B.,
RA Ho D.D.;
RT *HIV-1 subtypes, co-receptor usage, and CCR5 polymorphism.*;
RL AIDS Res. Hum. Retroviruses 0:0-0(1997).
DR EMBL: AF011538; AAB65738.1; -
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam: PF00001; 7tm_1; 1
DR PRINTS: PR00237; GPCR_RHODOPS.
DR PROSITE: PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE: PS00262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT NON_TER 352 352
SQ SEQUENCE 352 AA; 40523 MW; 4513DB983A28ACH2 CRC64;

Query Match 83.2%; Score 933; DB 6; Length 352;
Best Local Similarity 96.2%; Pred. No. 3.8e-79;
Matches 177; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 MDYQSSPTIDINYTSPEQCKINVKQIAARLLPPLYSLVIFGFGVGNMLVILLINCKR 60
DB 1 MDYQSSPTIDIDYTSPEQCKINVKQIAARLLPPLYSLVIFGFGVGNMLVILLINCKR 60
QY 61 LKSMTDIYLLNLAIISDLFLLTPFWAHYAAQWDFGNTMCLLTGLYFGFGSIFII 120
DB 61 LKSMTDIYLLNLAIISDLFLLTPFWAHYAAQWDFGNTMCLLTGLYFGFGSIFII 120
QY 121 LTTIDRYLAVHAFALKARTVTFGVTSVITWVAVFASLPFGIIFTRSQEGLHYTCSS 180
DB 121 LTTIDRYLAVHAFALKARTVTFGVTSVITWVAVFASLPFGIIFTRSQEGLHYTCSS 180
QY 181 HFPY 184
DB 181 HFPY 184

Search completed: June 3, 2003, 19:23:53
Job time : 44.5732 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2003, 19:18:46 ; Search time 12.2358 Seconds
(without alignments)
728.798 Million cell updates/sec

Title: US-09-938-703-6

Perfect score: 1122

Sequence: 1 MDYQVSSPIYDINTYSEPC.....AACGHLLGNKNSASVSK 215

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------|
| 1 | 958 | 85.4 | 352 | 1 | CCR5_HUMAN |
| 2 | 952 | 84.8 | 352 | 1 | CCR5_PANTR |
| 3 | 947 | 84.4 | 352 | 1 | CCR5_PORPY |
| 4 | 942 | 84.0 | 352 | 1 | CCR5_GORGO |
| 5 | 938 | 83.6 | 352 | 1 | CCR5_PYGBI |
| 6 | 938 | 83.6 | 352 | 1 | CCR5_PYGNE |
| 7 | 938 | 83.6 | 352 | 1 | CCR5_TRAFR |
| 8 | 936 | 83.6 | 352 | 1 | CCR5_TRAPH |
| 9 | 936 | 83.4 | 352 | 1 | CCR5_EVLLE |
| 10 | 933 | 83.2 | 352 | 1 | CCR5_MACMU |
| 11 | 933 | 83.2 | 352 | 1 | CCR5_PAPHA |
| 12 | 928 | 82.7 | 352 | 1 | CCR5_CERTO |
| 13 | 916 | 81.6 | 352 | 1 | CCR5_CERAE |
| 14 | 786 | 70.1 | 354 | 1 | CCR5_MOUSE |
| 15 | 775 | 69.1 | 354 | 1 | CCR5_RAT |
| 16 | 715.5 | 63.8 | 373 | 1 | CCR2_MOUSE |
| 17 | 714.5 | 63.7 | 373 | 1 | CCR2_RAT |
| 18 | 694 | 61.9 | 374 | 1 | CCR2_HUMAN |
| 19 | 688 | 61.3 | 360 | 1 | CCR2_MACMU |
| 20 | 598 | 53.3 | 355 | 1 | CCR1_HUMAN |
| 21 | 570.5 | 50.8 | 355 | 1 | CCR1_MOUSE |
| 22 | 562 | 50.1 | 355 | 1 | CCR1_MACMU |
| 23 | 547.5 | 48.8 | 360 | 1 | CCR4_HUMAN |
| 24 | 546.5 | 48.7 | 360 | 1 | CCR4_MOUSE |
| 25 | 535.5 | 47.7 | 359 | 1 | CCR3_MOUSE |
| 26 | 535.5 | 47.7 | 359 | 1 | CCR3_RAT |
| 27 | 516.5 | 46.0 | 358 | 1 | CCR3_CAVPO |
| 28 | 506.5 | 45.1 | 355 | 1 | CCR3_HUMAN |
| 29 | 499.5 | 44.5 | 355 | 1 | CCR3_MACMU |
| 30 | 489.5 | 43.6 | 355 | 1 | CCR3_CERAE |
| 31 | 489.5 | 43.6 | 356 | 1 | CCR3_MOUSE |
| 32 | 431 | 38.4 | 353 | 1 | CCR8_MOUSE |
| 33 | 427.5 | 38.1 | 355 | 1 | CCR8_HUMAN |

34 424 37.8 356 1 CCR8_MACMU
35 386 34.4 354 1 C3X1_RAT
36 383 34.1 354 1 C3X1_MOUSE
37 370 33.0 369 1 CCR9_MOUSE
38 365 32.5 355 1 C3X1_HUMAN
39 363 32.4 357 1 CCR9_HUMAN
40 345.5 30.8 378 1 CCR7_HUMAN
41 342.5 30.5 378 1 CCR7_MOUSE
42 334.5 29.8 352 1 CCR4_CERTO
43 331.5 29.5 352 1 CCR4_HUMAN
44 331.5 29.5 352 1 CCR4_MACMU
45 330.5 29.5 384 1 CCR6_HUMAN

ALIGNMENTS

RESULT 1
ID CCR5_HUMAN STANDARD; PRT: 352 AA.
AC P51681; O14692; O14693; O14695; O14696; O14697; O14698; O14699;
AC O14700; O14701; O14702; O14703; O14704; O14705; O14706; O14707;
AC O14708; O15538; Q9UPA4;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE C-C chemokine receptor type 5 (C-C CKR-5) (CC-CKR-5) (CCR-5) (CCR5)
DE (HIV-1 fusion co-receptor) (CHEMR13) (CD195 antigen).
GN CCR5 OR CCR5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID:9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-96241590; PubMed-8639485;
RA Samson M., Labbe O., Mollereau C., Vassart G., Parmentier M.;
RT "Molecular cloning and functional expression of a new human
CC-chemokine receptor gene."
RL Biochemistry 35:3362-3367(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-96291862; PubMed-8663314;
RA Raport C.J., Gosling J., Schweichart V.L., Gray P.W., Charo I.F.;
RT "Molecular cloning and functional characterization of a novel human
CC chemokine receptor (CCR5) for RANTES, MIP-beta, and MIP-lambda."
RL J. Biol. Chem. 271:17161-17166(1996).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE-96295970; PubMed-8699119;
RA Combadere C., Ahuja S.K., Tiffany H.L., Murphy P.M.;
RT "Cloning and functional expression of CC CCR5, a human monocyte CC
chemokine receptor selective for MIP-1(alpha), MIP-1(beta), and
RANTES."
RL J. Leukoc. Biol. 60:147-152(1996).
RN [4]
RP SEQUENCE FROM N.A.
RA McCombie W.R., Wilson R., Chen E., Gibbs R., Zuo L., Johnson D.,
Nhan M., Farnell L., Dedhia N., Ansari A., Mardis E., Schutz K.,
Gao J., la Bastide M., Kaplan N., Greco T., Touchman J.,
Muzny D., Chen C.N., Evans C., Fitzgerald M., See L.H., Tang M.,
Porcel B.M., Dragan Y., Giacalone J., Pae A., Powell E.,
Solinsky K.A., Desilva U., Diaz-Perez S., Zhou X., Yu Y.,
Watanabe M., Doggett N., Garcia D., Sagripanti J.L.;
RT Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE-98001387; PubMed-9343222;
RA Kuhnmann S.E., Platt E.J., Kozak S.L., Kabat D.;
RT "Polymorphisms in the CCR5 genes of African green monkeys and mice
implicate specific amino acids in infections by simian and human
immunodeficiency viruses."
RL J. Virol. 71:8642-8656(1997).

RN RP SEQUENCE FROM N.A., AND POLYMORPHISMS.
 RX MEDLINE-98022612; PubMed-9359654;
 RA Zhang L., Carruthers C.D., He T., Huang Y., Cao Y., Wang G., Hahn B.,
 RA Ho D.D.;
 RT "HIV type 1 subtypes, coreceptor usage, and CCR5 polymorphism.";
 RL AIDS Res. Hum. Retroviruses 13:1357-1366(1997).
 RN RP SEQUENCE FROM N.A.
 RX MEDLINE-98049523; PubMed-9388201;
 RA Mummidi S., Ahuja S.S., McDaniel B.L., Ahuja S.K.;
 RT "The human CC chemokine receptor 5 (CCR5) gene. Multiple transcripts
 with 5'-end heterogeneity, dual promoter usage, and evidence for
 polymorphisms within the regulatory regions and noncoding exons.";
 RL J. Biol. Chem. 272:30662-30671(1997).
 RN RP SEQUENCE FROM N.A., AND VARIANT ARG-178.
 RA Magierowska M., Barre-Sinoussi F., Issafras H., Theodorou I.,
 RA Debre P.;
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 RN RP CHARACTERIZATION OF ITS HIV-1 RECEPTOR FUNCTION.
 RX MEDLINE-96260017; PubMed-8649511;
 RA Deng H., Liu R., Elmeier W., Choe S., Unutmaz D., Burkhardt M.,
 RA di Marzio P., Marmon S., Sutton R.E., Hill C.M., Davis C.B.,
 RA Peiper S.C., Schall T.J., Littman D.R., Landau N.R.;
 RT "Identification of a major co-receptor for primary isolates of
 RT HIV-1.";
 RL Nature 381:661-666(1996).
 RN RP CHARACTERIZATION OF ITS HIV-1 RECEPTOR FUNCTION.
 RX MEDLINE-96260018; PubMed-8649512;
 RA Dragic T., Litwin V., Allaway G.P., Martin S.R., Huang Y.,
 RA Nagashima K.A., Cayanan C., Maddon P.J., Koup R.A., Moore J.P.,
 RA Paxton W.A.;
 RT "HIV-1 entry into CD4+ cells is mediated by the chemokine receptor
 RT CC-CR5.";
 RL Nature 381:667-673(1996).
 RN RP Sulfation.
 RX MEDLINE-99189752; PubMed-10089882;
 RA Farzan M., Mirzabekov T., Kolchinsky P., Wyatt R., Cayabyab M.,
 RA Gerard N.P., Gerard C., Sodroski J., Choe H.;
 RT "Tyrosine sulfation of the amino terminus of CCR5 facilitates HIV-1
 RT entry.";
 RL Cell 96:667-676(1999).
 CC -1- FUNCTION: RECEPTOR FOR A C-C TYPE CHEMOKINE. BINDS TO MIP-1-ALPHA,
 MIP-1-BETA AND RANTES AND SUBSEQUENTLY TRANSDUCES A SIGNAL BY
 INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL. MAY PLAY A ROLE
 IN THE CONTROL OF GRANULOCYTIC LINEAGE PROLIFERATION OR
 DIFFERENTIATION. ACTS AS CO-RECEPTOR WITH CD4 FOR PRIMARY NON-
 SYNCTYON-INDUCING STRAINS (NSI) (MACROPHAGE-TROPIC) OF HIV-1
 VIRUS. IT PROMOTES ENV-MEDIATED FUSION OF THE VIRUS.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: FOUND IN PROMYELOCYTIC CELLS.
 CC -1- PTM: SULFATION CONTRIBUTES TO THE EFFICIENCY OF HIV-1 ENTRY.
 CC -1- PTM: MODIFIED BY O-LINKED GLYCOSYLATION, BUT NOT BY N-LINKED
 GLYCOSYLATION.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch)

 DR EMBL; X91492; CAA62796.1; -
 DR EMBL; U54994; AAC50598.1; -
 DR EMBL; U57840; AAB17071.1; -
 DR EMBL; U95626; AAB57793.1; -
 DR EMBL; U83326; AAC51797.1; -
 DR EMBL; AF011500; AAB65700.1; -
 DR EMBL; AF011501; AAB65701.1; -
 DR EMBL; AF011502; AAB65702.1; -
 DR EMBL; AF011503; AAB65703.1; -
 DR EMBL; AF011505; AAB65705.1; -
 DR EMBL; AF011506; AAB65706.1; -
 DR EMBL; AF011507; AAB65707.1; -
 DR EMBL; AF011508; AAB65708.1; -
 DR EMBL; AF011509; AAB65709.1; -
 DR EMBL; AF011510; AAB65710.1; -
 DR EMBL; AF011511; AAB65711.1; -
 DR EMBL; AF011512; AAB65712.1; -
 DR EMBL; AF011513; AAB65713.1; -
 DR EMBL; AF011514; AAB65714.1; -
 DR EMBL; AF011515; AAB65715.1; -
 DR EMBL; AF011516; AAB65716.1; -
 DR EMBL; AF011517; AAB65717.1; -
 DR EMBL; AF011518; AAB65718.1; -
 DR EMBL; AF011519; AAB65719.1; -
 DR EMBL; AF011520; AAB65720.1; -
 DR EMBL; AF011521; AAB65721.1; -
 DR EMBL; AF011522; AAB65722.1; -
 DR EMBL; AF011523; AAB65723.1; -
 DR EMBL; AF011524; AAB65724.1; -
 DR EMBL; AF011525; AAB65725.1; -
 DR EMBL; AF011526; AAB65726.1; -
 DR EMBL; AF011527; AAB65727.1; -
 DR EMBL; AF011528; AAB65728.1; -
 DR EMBL; AF011529; AAB65729.1; -
 DR EMBL; AF011530; AAB65730.1; -
 DR EMBL; AF011531; AAB65731.1; -
 DR EMBL; AF011532; AAB65732.1; -
 DR EMBL; AF011533; AAB65733.1; -
 DR EMBL; AF011534; AAB65734.1; -
 DR EMBL; AF011535; AAB65735.1; -
 DR EMBL; AF011536; AAB65736.1; -
 DR EMBL; AF011537; AAB65737.1; -
 DR EMBL; AF031237; AAB94735.1; -
 DR EMBL; AF052539; AAD18131.1; -
 DR Genew; HGNC:1606; CCR5.
 DR MIM; 601373; -
 DR InterPro; IPR000276; GPCR_Rhodpsn.
 DR Pfam; PF00001; 7tm_1; 1.
 DR PRINTS; PR00237; GPCRHHODPSN.
 DR PROSITE; PS00237; G-PROTEIN RECEPTOR FL1; 1.
 DR PROSITE; PS00262; G-PROTEIN RECEPTOR FL2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation;
 KW Polymorphism.
 FT DOMAIN 1 30 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 31 58 1 (POTENTIAL).
 FT DOMAIN 59 68 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 69 89 2 (POTENTIAL).
 FT DOMAIN 90 102 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 103 124 3 (POTENTIAL).
 FT DOMAIN 125 141 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 142 166 4 (POTENTIAL).
 FT DOMAIN 167 198 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 199 218 5 (POTENTIAL).
 FT DOMAIN 219 235 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 236 260 6 (POTENTIAL).
 FT DOMAIN 261 277 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 278 301 7 (POTENTIAL).
 FT DOMAIN 302 352 CYTOPLASMIC (POTENTIAL).
 FT DISULFID 101 178 BY SIMILARITY.
 FT MOD_RES 3 3 SULFATION.
 FT MOD_RES 10 10 SULFATION.
 FT MOD_RES 14 14 SULFATION.
 FT MOD_RES 15 15 SULFATION.
 FT VARIANT 10 10 Y -> D (IN INCCR5-71A).
 FT VARIANT 29 29 A -> S (IN DESNP:1800939).
 FT VARIANT 31 31 R -> H (IN INCCR5-72A).

```
Query Match      85.4%; Score 958; DB 1; Length 352;
Best Local Similarity 100.0%; Pred. No. 2.3e-53;
Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDYQVSPYIDNYITSPCKINVKQIAARLLPPLSLVFIKFGVGNMVLILINCKR 60
DB 1 MDYQVSPYIDNYITSPCKINVKQIAARLLPPLSLVFIKFGVGNMVLILINCKR 60
QY 61 LKSMTDIYLNLAIISDLFFLTVPFWAHYAAQWDFGNTMCQLLTGLYIFGFSIGFII 120
DB 61 LKSMTDIYLNLAIISDLFFLTVPFWAHYAAQWDFGNTMCQLLTGLYIFGFSIGFII 120
QY 121 LITIDRYLAVHVAFAKARTVFGVVTSTVITWVAVFASLPGIIFTRSQEGLHYTCSS 180
DB 121 LITIDRYLAVHVAFAKARTVFGVVTSTVITWVAVFASLPGIIFTRSQEGLHYTCSS 180
QY 181 HPPY 184
DB 181 HPPY 184

RESULT 2
CCRS_PANTR STANDARD; PRT; 352 AA.
ID CCR5_PANTR
AC P56440; 002778;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE C-C chemokine receptor type 5 (C-CR-5) (CCR-5) (CCR5).
GN CCR5 OR CMKR5.
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX NCBI_TaxID=9598;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=97268687; PubMed=9108095;
RA Edinger A.L., Amedee A., Miller K., Doranz B.J., Endres M.,
RA Sharon M., Samson M., Lu Z.H., Clements J.E., Murphy-Corb M.,
RA Peliper S.C., Parmentier M., Broder C.C., Doms R.W.;
RT "Differential utilization of CCR5 by macrophage and T cell tropic
RT simian immunodeficiency virus strains.";
RL Proc. Natl. Acad. Sci. U.S.A. 94:4005-4010(1997).
RN [2]
SEQUENCE FROM N.A.
RA Zimmerman P.A., Buckler-White A., Alkhatib G.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
SEQUENCE FROM N.A.
RX MEDLINE=98022612; PubMed=9359654;
RA Zhang L., Carruthers C.D., He T., Huang Y., Cao Y., Wang G., Hahn B.,
RA Ho D.D.;
RT "HIV type 1 subtypes, coreceptor usage, and CCR5 polymorphism.";
RL AIDS Res. Hum. Retroviruses 13:1357-1366(1997).
RN [4]
SEQUENCE FROM N.A.
RX MEDLINE=97426118; PubMed=9282822;
RA Zacharova V., Zachar V., Goustin A.S.;
RT "Sequence of chemokine receptor gene CCR5 in chimpanzees, a natural
RT HIV type 1 host.";
RL AIDS Res. Hum. Retroviruses 13:1159-1161(1997).
RN [5]
SEQUENCE FROM N.A.
RX MEDLINE=98090115; PubMed=9430250;
RA Pretet J.-L., Zerbib A., Girard M., Guillet J.-G., Butor C.;
RT "Chimpanzee CXCR4 and CCR5 act as coreceptors for HIV type 1.";
RN AIDS Res. Hum. Retroviruses 13:1583-1587(1997).
CC -1- FUNCTION: RECEPTOR FOR A C-C TYPE CHEMOKINE. BINDS TO MIP-1-ALPHA,
CC MIP-1-BETA AND RANTES AND SUBSEQUENTLY TRANSDUCES A SIGNAL BY
CC INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL. MAY PLAY A ROLE
CC IN THE CONTROL OF GRANULOCYTIC LINEAGE PROLIFERATION OR
CC DIFFERENTIATION.
```

```
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL; AF005663; AAB62557.1; -
CC EMBL; U94329; AAB58446.1; -
CC EMBL; AF011542; AAB65742.1; -
CC EMBL; U97666; AAC51670.1; -
CC EMBL; AF011540; AAB65740.1; -
CC EMBL; U89797; AAC03717.1; -
CC InterPro; IPR000276; GPCR_Rhodopsn.
CC Pfam; PF00001; 7tm_1; 1.
CC PRINTS; PR00237; GPCRHOPOPSN.
CC PROSITE; PS00237; G_PROTEIN_RECP_F1_1; 1.
CC PROSITE; PS00262; G_PROTEIN_RECP_F1_2; 1.
CC KW G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation.
FT DOMAIN 1 30
FT TRANSMEM 31 58
FT DOMAIN 59 68
FT TRANSMEM 69 89
FT DOMAIN 90 102
FT TRANSMEM 103 124
FT DOMAIN 125 141
FT TRANSMEM 142 166
FT DOMAIN 167 198
FT TRANSMEM 199 218
FT DOMAIN 219 235
FT TRANSMEM 236 260
FT DOMAIN 261 277
FT TRANSMEM 278 301
FT DOMAIN 302 352
FT DISULFD 101 178
FT MOD_RES 3 3
FT MOD_RES 10 10
FT MOD_RES 14 14
FT MOD_RES 15 15
FT CARBOHYD 268 268
FT CONFLICT 123 123
FT CONFLICT 123 123
SQ SEQUENCE 352 AA; 40539 MW; 4A33E698B80FE34C CRC64;
Query Match      84.8%; Score 952; DB 1; Length 352;
Best Local Similarity 98.9%; Pred. No. 5.4e-53;
Matches 182; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDYQVSPYIDNYITSPCKINVKQIAARLLPPLSLVFIKFGVGNMVLILINCKR 60
DB 1 MDYQVSPYIDNYITSPCKINVKQIAARLLPPLSLVFIKFGVGNMVLILINCKR 60
QY 61 LKSMTDIYLNLAIISDLFFLTVPFWAHYAAQWDFGNTMCQLLTGLYIFGFSIGFII 120
DB 61 LKSMTDIYLNLAIISDLFFLTVPFWAHYAAQWDFGNTMCQLLTGLYIFGFSIGFII 120
QY 121 LITIDRYLAVHVAFAKARTVFGVVTSTVITWVAVFASLPGIIFTRSQEGLHYTCSS 180
DB 121 LITIDRYLAVHVAFAKARTVFGVVTSTVITWVAVFASLPGIIFTRSQEGLHYTCSS 180
QY 181 HPPY 184
DB 181 HPPY 184

RESULT 3
CCRS_PONPY STANDARD; PRT; 352 AA.
ID CCR5_PONPY
AC O97881;
DT 30-MAY-2000 (Rel. 39, Created)
```


DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE C-C chemokine receptor type 5 (C-CR-5) (CC-CR-5) (CCR-5) (CCR5).
 GN CCR5 OR CCR5.
 OS Pongo pygmaeus (Orangutan).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.
 OX NCBI_TaxID=9600;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99416438; PubMed=10486970;
 RA Zhang Y.-W., Ryder O.A., Zhang Y.-P.;
 RT "Sequence evolution of the CCR5 chemokine receptor gene in primates.";
 RL Mol. Biol. Evol. 16:1145-1154(1999).
 CC -1- FUNCTION: RECEPTOR FOR A C-C TYPE CHEMOKINE. BINDS TO MIP-1-ALPHA,
 CC MIP-1-BETA AND RANTES AND SUBSEQUENTLY TRANSDUCES A SIGNAL BY
 CC INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL. MAY PLAY A ROLE
 CC IN THE CONTROL OF GRANULOCYTIC LINEAGE PROLIFERATION OR
 CC DIFFERENTIATION.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF075446; AAD19858.1; -
 DR InterPro: IPR000276; GPCR_Rhodpsn.
 DR Pfam: PF00001; 7tm1.1;
 DR PRINTS: PR00237; GPCRHHODPSN.
 DR PROSITE: PS00237; G-PROTEIN RECEPTOR FL1; 1.
 DR PROSITE: PS00262; G-PROTEIN RECEPTOR FL2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation.
 FT DOMAIN 1 30 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 31 58 1 (POTENTIAL).
 FT DOMAIN 59 68 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 69 89 2 (POTENTIAL).
 FT DOMAIN 90 102 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 103 124 3 (POTENTIAL).
 FT DOMAIN 125 141 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 142 166 4 (POTENTIAL).
 FT DOMAIN 167 198 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 199 218 5 (POTENTIAL).
 FT DOMAIN 219 235 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 236 260 6 (POTENTIAL).
 FT DOMAIN 261 277 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 278 301 7 (POTENTIAL).
 FT DOMAIN 302 352 CYTOPLASMIC (POTENTIAL).
 FT DISULFID 101 178 BY SIMILARITY.
 FT MOD_RES 3 3 SULFATION (BY SIMILARITY).
 FT MOD_RES 10 10 SULFATION (BY SIMILARITY).
 FT MOD_RES 14 14 SULFATION (BY SIMILARITY).
 FT MOD_RES 15 15 SULFATION (BY SIMILARITY).
 SQ SEQUENCE 352 AA; 40527 MW; F4E2F47135AF658A CRC64;
 Query Match 84.4%; Score 947; DB 1; Length 352;
 Best Local Similarity 98.4%; Pred. No. 1.le-52;
 Matches 181; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MDYQVSSPIYDINVTSEPCQKINVKQIAARLLPPLSLVIFGFGVGNMVLILLNCKR 60
 DB 1 MDYQVSSPIYDINVTSEPCQKINVKQIAARLLPPLSLVIFGFGVGNMVLILLNCKR 60
 QY 61 LKSNWDIYLNLAIISDLFFLLTPVFWAHYAAQDFGNTMCQLLGLYFGFGSGLIFFII 120
 DB 61 LKSNWDIYLNLAIISDLFFLLTPVFWAHYAAQDFGNTMCQLLGLYFGFGSGLIFFII 120
 QY 121 LLTIDRYLAVHVAFAKARTVTFGVVTSVITWWVAVFASLPGIIFTRSQEGLHYTCSS 180
 DB 121 LLTIDRYLAVHVAFAKARTVTFGVVTSVITWWVAVFASLPGIIFTRSQEGLHYTCSS 180

DB 121 LLTIDRYLAVHVAFAKARTVTFGVVTSVITWWVAVFASLPGIIFTRSQEGLHYTCSS 180
 QY 181 HFPY 184
 DB 181 HFPY 184
 RESULT 4
 CCR5 GORGO STANDARD; PRT; 352 AA.
 ID P56439;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE C-C chemokine receptor type 5 (C-CR-5) (CC-CR-5) (CCR-5) (CCR5).
 GN CCR5 OR CCR5.
 OS Gorilla gorilla gorilla (Lowland gorilla).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
 OX NCBI_TaxID=9595;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97268687; PubMed=9108095;
 RA Edinger A.L., Amedee A., Miller K., Doranz B.J., Endres M.,
 RA Sharron M., Samson M., Lu Z.H., Clements J.E., Murphy-Corb M.,
 RA Feiler S.C., Parmentier M., Broder C.C., Doms R.W.;
 RT "Differential utilization of CCR5 by macrophage and T cell tropic
 RT simian immunodeficiency virus strains.";
 RL Proc. Natl. Acad. Sci. U.S.A. 94:4005-4010(1997).
 CC -1- FUNCTION: RECEPTOR FOR A C-C TYPE CHEMOKINE. BINDS TO MIP-1-ALPHA,
 CC MIP-1-BETA AND RANTES AND SUBSEQUENTLY TRANSDUCES A SIGNAL BY
 CC INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL. MAY PLAY A ROLE
 CC IN THE CONTROL OF GRANULOCYTIC LINEAGE PROLIFERATION OR
 CC DIFFERENTIATION.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF005659; AAB62553.1; -
 DR InterPro: IPR000276; GPCR_Rhodpsn.
 DR Pfam: PF00001; 7tm1.1;
 DR PRINTS: PR00237; GPCRHHODPSN.
 DR PROSITE: PS00237; G-PROTEIN RECEPTOR FL1; 1.
 DR PROSITE: PS00262; G-PROTEIN RECEPTOR FL2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation.
 FT DOMAIN 1 30 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 31 58 1 (POTENTIAL).
 FT DOMAIN 59 68 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 69 89 2 (POTENTIAL).
 FT DOMAIN 90 102 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 103 124 3 (POTENTIAL).
 FT DOMAIN 125 141 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 142 166 4 (POTENTIAL).
 FT DOMAIN 167 198 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 199 218 5 (POTENTIAL).
 FT DOMAIN 219 235 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 236 260 6 (POTENTIAL).
 FT DOMAIN 261 277 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 278 301 7 (POTENTIAL).
 FT DISULFID 101 178 BY SIMILARITY.
 FT MOD_RES 3 3 SULFATION (BY SIMILARITY).
 FT MOD_RES 10 10 SULFATION (BY SIMILARITY).
 FT MOD_RES 14 14 SULFATION (BY SIMILARITY).
 FT MOD_RES 15 15 SULFATION (BY SIMILARITY).
 SQ SEQUENCE 352 AA; 40515 MW; D0E6FCB9F5EAC84 CRC64;

Query Match 84.0%; Score 942; DB 1; Length 352;
Best Local Similarity 97.8%; Pred. No. 2.3e-52;
Matches 180; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MDYQSSPIYDINTYSEPCQKINVKQIAARLLPPLYSIVFIFGVGNMVLILLINCKR 60
Dy 1 MDYQSSPTDIDYTTSEPCQKINVKQIAARLLPPLYSIVFIFGVGNMVLILLINCKR 60
Qy 61 LKSMTDIYLLNLAIISDLFFLLVFPWAHYAAQWDFGNTMCLLTGLYFIFGFSGIFPII 120
Dy 61 LKSMTDIYLLNLAIISDLFFLLVFPWAHYAAQWDFGNTMCLLTGLYFIFGFSGIFPII 120
Qy 121 LTTIDRYLAVHAFALKARTVFGVTSVITWVAVFASLPGIIFTRSQEGLHYTCSS 180
Dy 121 LTTIDRYLAVHAFALKARTVFGVTSVITWVAVFASLPGIIFTRSQEGLHYTCSS 180
Qy 181 HPFY 184
Dy 181 HPFY 184

RESULT 5
CKR5_PYGBI STANDARD; PRT; 352 AA.
AC 097860;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE C-C chemokine receptor type 5 (C-C CKR-5) (CCR-5) (CCR5).
GN CCR5 OR CMKR5.
OS Pygathrix bieti (Black snub-nosed monkey) (Rhinopithecus bieti).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Colobinae;
OC Pygathrix.
OX NCBI_TaxID=61621;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-99416438; PubMed-10486970;
RA Zhang Y.-W., Ryder O.A., Zhang Y.-P.;
RT "Sequence evolution of the CCR5 chemokine receptor gene in primates.";
RL Mol. Biol. Evol. 16:1145-1154(1999).
CC MIP-1-BETA AND RANTES AND SUBSEQUENTLY TRANSDUCES A SIGNAL BY
CC INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL. MAY PLAY A ROLE
CC IN THE CONTROL OF GRANULOCYTIC LINEAGE PROLIFERATION OR
CC DIFFERENTIATION.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; AF075445; AAD19857.1; -
CC InterPro; IPR000276; GPCR_Rhodpsn.
CC Pfam; PF00001; 7tm1.1;
CC PRINTS; PR00237; GPCRHOOPS.
CC PROSITE; PS00237; G-PROTEIN_RECP_F1_1; 1.
CC PROSITE; PS00262; G-PROTEIN_RECP_F1_2; 1.
CC G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation.
KW DOMAIN 1 30 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 31 58 1 (POTENTIAL).
FT TRANSMEM 59 68 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 69 89 2 (POTENTIAL).
FT TRANSMEM 90 102 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 103 124 3 (POTENTIAL).
FT TRANSMEM 125 141 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 142 166 4 (POTENTIAL).
FT TRANSMEM

FT DOMAIN 167 198 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 199 218 5 (POTENTIAL).
FT DOMAIN 219 235 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 236 260 6 (POTENTIAL).
FT DOMAIN 261 277 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 278 301 7 (POTENTIAL).
FT DOMAIN 302 352 CYTOPLASMIC (POTENTIAL).
FT DISULFID 101 178 BY SIMILARITY.
FT MOD_RES 3 3 SULFATION (BY SIMILARITY).
FT MOD_RES 10 10 SULFATION (BY SIMILARITY).
FT MOD_RES 14 14 SULFATION (BY SIMILARITY).
FT MOD_RES 15 15 SULFATION (BY SIMILARITY).
SQ SEQUENCE 352 AA; 4366F142730F938F CRC64;
Query Match 83.8%; Score 938; DB 1; Length 352;
Best Local Similarity 96.2%; Pred. No. 4e-52;
Matches 177; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MDYQSSPIYDINTYSEPCQKINVKQIAARLLPPLYSIVFIFGVGNMVLILLINCKR 60
Dy 1 MDYQSSPTDIDYTTSEPCQKINVKQIAARLLPPLYSIVFIFGVGNMVLILLINCKR 60
Qy 61 LKSMTDIYLLNLAIISDLFFLLVFPWAHYAAQWDFGNTMCLLTGLYFIFGFSGIFPII 120
Dy 61 LKSMTDIYLLNLAIISDLFFLLVFPWAHYAAQWDFGNTMCLLTGLYFIFGFSGIFPII 120
Qy 121 LTTIDRYLAVHAFALKARTVFGVTSVITWVAVFASLPGIIFTRSQEGLHYTCSS 180
Dy 121 LTTIDRYLAVHAFALKARTVFGVTSVITWVAVFASLPGIIFTRSQEGLHYTCSS 180
Qy 181 HPFY 184
Dy 181 HPFY 184

RESULT 6
CKR5_PYGNE STANDARD; PRT; 352 AA.
AC 097882;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE C-C chemokine receptor type 5 (C-C CKR-5) (CCR-5) (CCR5).
GN CCR5 OR CMKR5.
OS Pygathrix nemaeus (Dove langur).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Colobinae;
OC Pygathrix.
OX NCBI_TaxID=54133;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-99416438; PubMed-10486970;
RA Zhang Y.-W., Ryder O.A., Zhang Y.-P.;
RT "Sequence evolution of the CCR5 chemokine receptor gene in primates.";
RL Mol. Biol. Evol. 16:1145-1154(1999).
CC MIP-1-BETA AND RANTES AND SUBSEQUENTLY TRANSDUCES A SIGNAL BY
CC INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL. MAY PLAY A ROLE
CC IN THE CONTROL OF GRANULOCYTIC LINEAGE PROLIFERATION OR
CC DIFFERENTIATION.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; AF075448; AAD19860.1; -
CC InterPro; IPR000276; GPCR_Rhodpsn.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC
CC THE SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AF075442; AAD19854.1; -
CC InterPro: IPR000276; GPCR_Rhodpsn.
CC Pfam: PF00001; 7tm_1.
CC DR PRINIS: PR00237; GPCRHOODPSN.
CC DR PROSITE: PS00237; G-PROTEIN RECP_FL1; 1.
CC DR PROSITE: PS0262; G-PROTEIN RECP_FL2; 1.
CC KW G-protein coupled receptor; transmembrane; Glycoprotein; Sulfation.
CC FT DOMAIN 1 30 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 31 58 1 (POTENTIAL).
CC FT DOMAIN 59 68 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 69 89 2 (POTENTIAL).
CC FT DOMAIN 90 102 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 103 124 3 (POTENTIAL).
CC FT DOMAIN 125 141 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 142 166 4 (POTENTIAL).
CC FT DOMAIN 167 198 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 199 218 5 (POTENTIAL).
CC FT DOMAIN 219 235 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 236 260 6 (POTENTIAL).
CC FT DOMAIN 261 277 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 278 301 7 (POTENTIAL).
CC FT DOMAIN 302 352 CYTOPLASMIC (POTENTIAL).
CC FT DISULFID 101 178 BY SIMILARITY.
CC FT MOD_RES 3 3 SULFATION (BY SIMILARITY).
CC FT MOD_RES 10 10 SULFATION (BY SIMILARITY).
CC FT MOD_RES 14 14 SULFATION (BY SIMILARITY).
CC FT MOD_RES 15 15 SULFATION (BY SIMILARITY).
CC SQ SEQUENCE 352 AA; 40496 MW; 4366F148C25938F CRC64;

CC Query Match 93.6%; Score 938; DB 1; Length 352;
CC Best Local Similarity 86.2%; Pred. No. 4e-52;
CC Matches 177; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

CC QY 1 MDYVSSPIIDINYYSEPCQKINVKQIAARLLPPLYSIFVFGVGMVLVILLINCKR 60
CC DB 1 MDYVSSPTIDIDYITSEPCQKVNKQIAARLLPPLYSIFVFGVGNLVLLINCKR 60
CC QY 61 LKSMTDIVLLMALISDLFLLPVFWAHYAAQWDFGNTMCQLLTGLYFIFGFSIGFIIF 120
CC DB 61 LKSMTDIVLLMALISDLFLLPVFWAHYAAQWDFGNTMCQLLTGLYFIFGFSIGFIIF 120
CC QY 121 LFTIDRYLAVHVAFFALKARTVTEGWTIVSTVWVAVFASLPGLIFTRSQKGLHYTCSS 180
CC DB 121 LFTIDRYLAVHVAFFALKARTVTEGWTIVSTVWVAVFASLPGLIFTRSQKGLHYTCSS 180
CC QY 181 HFPY 184
CC DB 181 HFPY 184

CC RESULT 8
CC CKR5_TRAPH
CC ID CKR5_TRAPH STANDARD; PRT: 352 AA.
CC AC O97879;
CC DT 30-MAY-2000 (Rel. 39, Created)
CC DT 30-MAY-2000 (Rel. 39, Last sequence update)
CC DT 15-JUN-2002 (Rel. 41, Last annotation update)
CC DE C-C chemokine receptor type 5 (C-C CKR-5) (CCR-5) (CCR5).
CC GN CCR5 OR CKR5.
CC OS Trachypithecus playrei (Playre's leaf monkey).
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae; Colobinae;

OC Trachypithecus.
 OX NCBI_TaxID=61618;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99416438; PubMed=10486970;
 RA Zhang Y.-W., Ryder O.A., Zhang Y.-P.;
 RT "Sequence evolution of the CCR5 chemokine receptor gene in primates.";
 RL Mol. Biol. Evol. 16:1145-1154(1999).
 CC -1- FUNCTION: RECEPTOR FOR A C-C TYPE CHEMOKINE. BINDS TO MIP-1-ALPHA,
 CC MIP-1-BETA AND RANTES AND SUBSEQUENTLY TRANSDUCES A SIGNAL BY
 CC INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL. MAY PLAY A ROLE
 CC IN THE CONTROL OF GRANULOCYTIC LINEAGE PROLIFERATION OR
 CC DIFFERENTIATION.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF075443; AD19855.1;
 CC InterPro: IPR000276; GPCR_Rhodpsn.
 CC Pfam: PF00001; 7tm.1; 1.
 CC PRINTS: PR00237; GPCRHHODPSN.
 CC PROSITE: PS00237; G-PROTEIN RECF_FL_1; 1.
 CC PROSITE: PS0262; G-PROTEIN RECF_FL_2; 1.
 CC G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation.
 KW DOMAIN 1 30
 FT TRANSMEM 31 58
 FT DOMAIN 59 68
 FT TRANSMEM 69 89
 FT DOMAIN 90 102
 FT TRANSMEM 103 124
 FT DOMAIN 125 141
 FT TRANSMEM 142 166
 FT DOMAIN 167 198
 FT TRANSMEM 199 218
 FT TRANSMEM 219 235
 FT DOMAIN 236 260
 FT TRANSMEM 261 277
 FT TRANSMEM 278 301
 FT DOMAIN 302 352
 FT DISULFID 101 178
 FT MOD_RES 3 3
 FT MOD_RES 10 10
 FT MOD_RES 14 14
 FT MOD_RES 15 15
 FT MOD_RES 15 15
 SQ SEQUENCE 352 AA; 40509 MW; 4366f148d3a5938f CRC64;
 Query Match 83.6%; Score 938; DB 1; Length 352;
 Best Local Similarity 96.2%; Pred. No. 4e-52;
 Matches 177; Conservative 6; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MDYQSSPYDINVTSEPCQKINVKQIAARLLPPLYSLVIFGFGVGNMVLILINCKR 60
 DB 1 MDYQSSPYDIDYITSEPCQKINVKQIAARLLPPLYSLVIFGFGVGNMVLILINCKR 60
 QY 61 LKSMTDIYLLNLAIISDLFLLVFPWAHYAAQWDRGNTWCQLLTGLYFGFGSIFITII 120
 DB 61 LKSMTDIYLLNLAIISDLFLLVFPWAHYAAQWDRGNTWCQLLTGLYFGFGSIFITII 120
 QY 121 LTRIDRLAVHVAFLKARTVTFGVTSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
 DB 121 LTRIDRLAVHVAFLKARTVTFGVTSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
 QY 181 HFPY 184
 DB 181 HFPY 184

RESULT 9
 ID CCR5_HYLLE STANDARD; PRT; 352 AA.
 AC 097883;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE C-C chemokine receptor type 5 (C-CR-5) (CCR-5) (CCR5).
 GN CCR5 OR CCR5R3.
 OS Hylobates leucogenys (White-cheeked gibbon).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hylobatidae; Hylobates.
 OX NCBI_TaxID=61853;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99416438; PubMed=10486970;
 RA Zhang Y.-W., Ryder O.A., Zhang Y.-P.;
 RT "Sequence evolution of the CCR5 chemokine receptor gene in primates.";
 RL Mol. Biol. Evol. 16:1145-1154(1999).
 CC -1- FUNCTION: RECEPTOR FOR A C-C TYPE CHEMOKINE. BINDS TO MIP-1-ALPHA,
 CC MIP-1-BETA AND RANTES AND SUBSEQUENTLY TRANSDUCES A SIGNAL BY
 CC INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL. MAY PLAY A ROLE
 CC IN THE CONTROL OF GRANULOCYTIC LINEAGE PROLIFERATION OR
 CC DIFFERENTIATION.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF075451; AD19863.1;
 CC InterPro: IPR000276; GPCR_Rhodpsn.
 CC Pfam: PF00001; 7tm.1; 1.
 CC PRINTS: PR00237; GPCRHHODPSN.
 CC PROSITE: PS00237; G-PROTEIN RECF_FL_1; 1.
 CC PROSITE: PS0262; G-PROTEIN RECF_FL_2; 1.
 CC G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation.
 KW DOMAIN 1 30
 FT TRANSMEM 31 58
 FT DOMAIN 59 68
 FT TRANSMEM 69 89
 FT DOMAIN 90 102
 FT TRANSMEM 103 124
 FT DOMAIN 125 141
 FT TRANSMEM 142 166
 FT DOMAIN 167 198
 FT TRANSMEM 199 218
 FT TRANSMEM 219 235
 FT TRANSMEM 236 260
 FT TRANSMEM 261 277
 FT TRANSMEM 278 301
 FT DOMAIN 302 352
 FT DISULFID 101 178
 FT MOD_RES 3 3
 FT MOD_RES 10 10
 FT MOD_RES 14 14
 FT MOD_RES 15 15
 FT MOD_RES 15 15
 SQ SEQUENCE 352 AA; 40445 MW; 4f8e4f344ce7c91 CRC64;
 Query Match 83.4%; Score 936; DB 1; Length 352;
 Best Local Similarity 97.3%; Pred. No. 5.4e-52;
 Matches 179; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 QY 1 MDYQSSPYDINVTSEPCQKINVKQIAARLLPPLYSLVIFGFGVGNMVLILINCKR 60
 DB 1 MDYQSSPYDIDYITSEPCQKINVKQIAARLLPPLYSLVIFGFGVGNMVLILINCKR 60
 QY 61 LKSMTDIYLLNLAIISDLFLLVFPWAHYAAQWDRGNTWCQLLTGLYFGFGSIFITII 120

```
Db 11 LKSMTDIYLNLAIASDLFFLLTPVFWAHYAAQWDFGNTMCQLLTGLYFGFGSIFPII 120
QY 121 LTTDRYLAVHVAFAKARTVTVGVVTSVITWVAVFASIPGLIIFTRSQEGLHYTCSS 180
Db 121 LTTDRYLAVHVAFAKARTVTVGVVTSVITWVAVFASIPGLIIFTRSQEGLHYTCSS 180
QY 181 HPFY 184
Db 181 HPFY 184

RESULT 10
CKR5_MACMU STANDARD; PRT; 352 AA.
AC P79436; 002746;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE C-C chemokine receptor type 5 (C-C CKR-5) (CC-CKR-5) (CCR-5) (CCR5).
GN CKR5 OR CKR5.
OS Macaca mulatta (Rhesus macaque),
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey), and
OS Macaca nemestrina (Pig-tailed macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_TaxID=9544, 9541, 9545;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=M.mulatta;
RX MEDLINE=97184592; PubMed=9032394;
RA Marcon L., Choe H., Martin K.A., Farzan M., Ponath P.D., Wu L.,
RA Newman W., Gerard N., Gerard C., Sodroski J.;
RT Utilization of C-C chemokine receptor 5 by the envelope
RT glycoproteins of a pathogenic simian immunodeficiency virus,
RT SIVmac239.
RT J. Virol. 71:2522-2527 (1997).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=M.mulatta;
RX MEDLINE=97213934; PubMed=9060623;
RA Chen Z., Zhou P., Ho D.D., Landau N.R., Marx P.A.;
RT "Genetically divergent strains of simian immunodeficiency virus use
RT CKR5 as a coreceptor for entry."
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=M.mulatta;
RX MEDLINE=21334176; PubMed=11461684;
RA Margulies B.J., Hauer D.A., Clements J.E.;
RT "Identification and comparison of eleven rhesus macaque chemokine
RT receptors."
RN [4]
RP AIDS Res. Hum. Retroviruses 17:981-986 (2001).
RN [5]
RP SEQUENCE FROM N.A.
RC SPECIES=M.mulatta;
RX MEDLINE=97268687; PubMed=9108095;
RA Edinger A.L., Amedee A., Miller K., Doranz B.J., Endres M.,
RA Sharon M., Samson M., Lu Z.H., Clements J.E., Murphy-Corb M.,
RA Peiper S.C., Parmentier M., Broder C.C., Doms R.W.;
RT Differential utilization of CKR5 by macrophage and T cell tropic
RT simian immunodeficiency virus strains.
RN [6]
RP Proc. Natl. Acad. Sci. U.S.A. 94:4005-4010 (1997).
CC -1- FUNCTION: RECEPTOR FOR A C-C TYPE CHEMOKINE. BINDS TO MIP-1-ALPHA,
CC MIP-1-BETA AND RANTES AND SUBSEQUENTLY TRANSDUCES A SIGNAL BY
CC INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL. MAY PLAY A ROLE
CC IN THE CONTROL OF GRANULOCYTIC LINEAGE PROLIFERATION OR
CC DIFFERENTIATION.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
```

```
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U77672; AAC511109.1; -
DR EMBL; U73739; AAC511158.1; -
DR EMBL; U96762; AAC34132.1; -
DR EMBL; AF005660; AB62554.1; -
DR EMBL; AF005661; AB62555.1; -
DR EMBL; AF005662; AB62556.1; -
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHHODPSN.
DR PROSITE; PS00237; G-PROTEIN_RECEP_F1.1; 1.
DR PROSITE; PS0262; G-PROTEIN_RECEP_F1.2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation.
FT DOMAIN 1 30 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 31 58 1 (POTENTIAL).
FT DOMAIN 59 68 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 69 89 2 (POTENTIAL).
FT DOMAIN 90 102 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 103 124 3 (POTENTIAL).
FT DOMAIN 125 141 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 142 166 4 (POTENTIAL).
FT DOMAIN 167 198 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 199 218 5 (POTENTIAL).
FT DOMAIN 219 235 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 236 260 6 (POTENTIAL).
FT DOMAIN 261 277 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 278 301 7 (POTENTIAL).
FT DOMAIN 302 352 CYTOPLASMIC (POTENTIAL).
FT DISULFID 101 178 BY SIMILARITY.
FT MOD_RES 3 3 SULFATION (BY SIMILARITY).
FT MOD_RES 10 10 SULFATION (BY SIMILARITY).
FT MOD_RES 14 14 SULFATION (BY SIMILARITY).
FT MOD_RES 15 15 SULFATION (BY SIMILARITY).
FT CARBOHYD 268 268 N-LINKED (GLCNAC...) (POTENTIAL).
FT CONFLICT 241 241 M -> I (IN REF. 3).
FT CONFLICT 292 292 I -> M (IN REF. 3).
SQ SEQUENCE 352 AA; 40507 MW; 58B96C85909FACB2 CRC64;

Query Match 83.2%; Score 933; DB 1; Length 352;
Best Local Similarity 96.2%; Pred. No. 8.3e-52;
Matches 177; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 MDYVSSPIYDINXYTSBPCKINVKQIAARLLPPLYSILVIFGVGNMVLILINCKR 60
Db 1 MDYVSSPIYDIDYTSBPCKINVKQIAARLLPPLYSILVIFGVGNMVLILINCKR 60
QY 61 LKSMTDIYLNLAIASDLFFLLTPVFWAHYAAQWDFGNTMCQLLTGLYFGFGSIFPII 120
Db 61 LKSMTDIYLNLAIASDLFFLLTPVFWAHYAAQWDFGNTMCQLLTGLYFGFGSIFPII 120
QY 121 LTTDRYLAVHVAFAKARTVTVGVVTSVITWVAVFASIPGLIIFTRSQEGLHYTCSS 180
Db 121 LTTDRYLAVHVAFAKARTVTVGVVTSVITWVAVFASIPGLIIFTRSQEGLHYTCSS 180
QY 181 HPFY 184
Db 181 HPFY 184

RESULT 11
CKR5_PAPHA STANDARD; PRT; 352 AA.
AC P56441;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE C-C chemokine receptor type 5 (C-C CKR-5) (CCR-5) (CCR5).
```

| | | | | |
|-----|----------|---------|-----------|-----------------------------------|
| FT | DISULFID | 101 | 178 | BY SIMILARITY. |
| FT | MOD_RES | 3 | 3 | SULFATION (BY SIMILARITY). |
| FT | MOD_RES | 10 | 10 | SULFATION (BY SIMILARITY). |
| FT | MOD_RES | 14 | 14 | SULFATION (BY SIMILARITY). |
| FT | MOD_RES | 15 | 15 | SULFATION (BY SIMILARITY). |
| FT | CARBOHYD | 268 | 268 | N-LINKED (GLCNAC. .) (POTENTIAL). |
| SEQ | SEQUENCE | 352 AA; | 40489 MW; | 5E1504A9BA1FE8B2 CRC64; |

Query Match 83.2%; Score 933; DB 1; Length 352;
 Best Local Similarity 96.2%; Pred. No. 8.3e-52;
 Matches 177; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

| | | | |
|----|-----|--|-----|
| QY | 1 | MDYQSSPTIDINXYTSEPCQKINVKQIARLLPPLSLVFIKFGVGNMLVILINCKR | 60 |
| DB | 1 | MDYQSSPTIDIDYIYSEPCQKINVKQIARLLPPLSLVFIKFGVGNMLVILINCKR | 60 |
| QY | 61 | LKSMTDIYLLMLAISDLFFLLVTFVFWAHTAAQWDFGNMCCQLTGLYFIKFGVSGIIFII | 120 |
| DB | 61 | LKSMTDIYLLMLAISDLFFLLVTFVFWAHTAAQWDFGNMCCQLTGLYFIKFGVSGIIFII | 120 |
| QY | 121 | LTTIDRYLVAVHAFKAKRTVFGVTSYITWVAVFASLPGLIIFTRSQEGLHYTCSS | 180 |
| DB | 121 | LTTIDRYLVAVHAFKAKRTVFGVTSYITWVAVFASLPGLIIFTRSQEGLHYTCSS | 180 |
| QY | 181 | HPFY 184 | |
| DB | 181 | HPFY 184 | |

RESULT 12
 CKR5_CERTO
 ID CKR5_CERTO STANDARD; PRT: 352 AA.
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DT C-C chemokine receptor type 5 (C-C CKR-5) (CCR-5) (CCR5).
 DE CKR5 OR CMK3R5.
 OS Cercopithecus torquatus atys (Red-crowned mangabey) (Sooty mangabey).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
 OC Cercopithecinae; Cercopithecus.
 OX NCBI_Taxid=9531;
 RN [1]
 RP STRAIN=Isolate 079, 085, 087, and 089;
 RC MEDLINE=98321155; PubMed=9656999;
 RA Chen Z., Gettie A., Ho D.D., Marx P.A.;
 RT "Primary SIVsm isolates use the CKR5 coreceptor from sooty mangabeys naturally infected in west Africa: a comparison of coreceptor usage of primary SIVsm, HIV-2, and SIVmac.";
 RL Virology 246:113-124(1998).
 CC -1- FUNCTION: RECEPTOR FOR A C-C TYPE CHEMOKINE. BINDS TO MIP-1-ALPHA, MIP-1-BETA AND RANTES AND SUBSEQUENTLY TRANSDUCES A SIGNAL BY INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL. MAY PLAY A ROLE IN THE CONTROL OF GRANULOCYTIC LINEAGE PROLIFERATION OR DIFFERENTIATION.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation in the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announcement> or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF051902; AAC39830.1; -
 CC EMBL; AF051903; AAC39831.1; -
 CC EMBL; AF051904; AAC39832.1; -
 CC EMBL; AF051905; AAC39833.1; -
 CC InterPro: IPR000276; GPCR_Rhodpsn.

DR PFAM: PF00001; 7tm.1; 1.
 DR PRINTS: PR00237; GPCRHOOPS.
 DR PROSITE: PS00237; G_PROTEIN_RECP_F1_1; 1.
 DR PROSITE: PS0262; G_PROTEIN_RECP_F1_2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation.
 FT DOMAIN 1 30
 FT TRANSMEM 31 58
 FT DOMAIN 59 68
 FT TRANSMEM 69 89
 FT DOMAIN 90 102
 FT TRANSMEM 103 124
 FT DOMAIN 125 141
 FT TRANSMEM 142 166
 FT DOMAIN 167 198
 FT TRANSMEM 199 219
 FT DOMAIN 219 235
 FT TRANSMEM 236 260
 FT DOMAIN 261 277
 FT TRANSMEM 278 301
 FT DOMAIN 302 325
 FT TRANSMEM 326 352
 FT DISULFID 101 178
 FT MOD_RES 3 3
 FT VARIANT 10 10
 FT MOD_RES 14 14
 FT MOD_RES 15 15
 FT VARIANT 2 2
 FT VARIANT 3 3
 FT VARIANT 25 25
 FT VARIANT 100 100
 FT VARIANT 107 107
 FT VARIANT 134 134
 FT VARIANT 146 146
 FT VARIANT 340 340
 SQ SEQUENCE 352 AA; 40489 MW; 20A196E2D47E43CA CRC64;
 Query Match 82.7%; Score 928; DB 1; Length 352;
 Best Local Similarity 95.1%; Pred. No. 1.7e-51;
 Matches 176; Conservative 5; Mismatches 3; Indels 0; Gaps 0;
 QY 1 MDYQVSPYIDYNTSEPCQKINVKQIAARLLPPLYSIVIFGFGVGNMLVILLINCKR 60
 DB 1 MDYQVSPYIDYNTSEPCQKINVKQIAARLLPPLYSIVIFGFGVGNMLVILLINCKR 60
 QY 61 LKSMTDIYLLNLALSIDFLFTVPFWAHYAAQWDFGNTWCQLLTGLYFIFGFSGIFII 120
 DB 61 LKSMTDIYLLNLALSIDFLFTVPFWAHYAAQWDFGNTWCQLLTGLYFIFGFSGIFII 120
 QY 121 LLTIDRYLAVHVAFAKARTVFGVTSVITWVAVFASLPGLIIFRSQKGLHYTCSS 180
 DB 121 LLTIDRYLAVHVAFAKARTVFGVTSVITWVAVFASLPGLIIFRSQKGLHYTCSS 180
 QY 181 HEPY 184
 DB 181 HEPY 184
 RESULT 13
 CKR5_CERAE
 ID CKR5_CERAE STANDARD; PRT; 352 AA.
 AC P56493;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE C-C chemokine receptor type 5 (C-C CKR-5) (CC-CKR-5) (CCR-5) (CCR5).
 GN CCR5 OR CMKR5.
 OS Cercopithecus aethiops (Green monkey) (Grivet).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecinae; Cercopithecus.
 OX NCBI_TaxID=9534;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;

RX MEDLINE=98001387; PubMed=9343222;
 RA Kuhlmann S.E., Platt E.J., Kozak S.L., Kabat D.;
 RT "Polymorphisms in the CCR5 genes of African green monkeys and mice
 RT implicate specific amino acids in infections by simian and human
 RT immunodeficiency viruses.";
 RL J. Virol. 71:8642-8656(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Murayama Y., Matsunaga S., Inoue-Murayama M.;
 RT "cDNA sequence of African green monkey CCR-5 chemokine receptor
 RT gene.";
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
 CC 1- FUNCTION: RECEPTOR FOR A C-C TYPE CHEMOKINE. BINDS TO MIP-1-ALPHA,
 CC MIP-1-BETA AND RANTES AND SUBSEQUENTLY TRANSDUCES A SIGNAL BY
 CC INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL. MAY PLAY A ROLE
 CC IN THE CONTROL OF GRANULOCYTIC LINEAGE PROLIFERATION OR
 CC DIFFERENTIATION.
 CC 1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC 1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: U83324; AAC51795.1; -;
 CC EMBL: U83325; AAC51796.1; -;
 CC EMBL: AB015944; BAA31328.1; -;
 CC InterPro: IPR000276; GPCR_Rhodpsn.
 CC Pfam: PF00001; 7tm.1; 1.
 CC PRINTS: PR00237; GPCRHOOPS.
 CC PROSITE: PS00237; G_PROTEIN_RECP_F1_1; 1.
 CC PROSITE: PS0262; G_PROTEIN_RECP_F1_2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation;
 KW Polymorphism.
 FT DOMAIN 1 30
 FT TRANSMEM 31 58
 FT DOMAIN 59 68
 FT TRANSMEM 69 89
 FT DOMAIN 90 102
 FT TRANSMEM 103 124
 FT DOMAIN 125 141
 FT TRANSMEM 142 166
 FT DOMAIN 167 198
 FT TRANSMEM 199 219
 FT DOMAIN 219 235
 FT TRANSMEM 236 260
 FT DOMAIN 261 277
 FT TRANSMEM 278 301
 FT DOMAIN 302 325
 FT DISULFID 101 178
 FT MOD_RES 3 3
 FT MOD_RES 10 10
 FT MOD_RES 15 15
 FT VARIANT 14 14
 FT VARIANT 352 352
 SQ SEQUENCE 352 AA; 40561 MW; 7F52E690C72EC29A CRC64;
 Query Match 81.6%; Score 916; DB 1; Length 352;
 Best Local Similarity 95.1%; Pred. No. 9.4e-51;
 Matches 175; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
 QY 1 MDYQVSPYIDYNTSEPCQKINVKQIAARLLPPLYSIVIFGFGVGNMLVILLINCKR 60
 DB 1 MDYQVSPYIDYNTSEPCQKINVKQIAARLLPPLYSIVIFGFGVGNMLVILLINCKR 60
 QY 61 LKSMTDIYLLNLALSIDFLFTVPFWAHYAAQWDFGNTWCQLLTGLYFIFGFSGIFII 120
 DB 61 LKSMTDIYLLNLALSIDFLFTVPFWAHYAAQWDFGNTWCQLLTGLYFIFGFSGIFII 120

QY 121 LFIIDRYLAVHVAFAKARTVFGVTVSVITWVAVFASLPGIIFRSQREGLHYTCSS 180
 DB 121 LFIIDRYLAVHVAFAKARTVFGVTVSVITWVAVFASLPGIIFRSQREGLHYTCSS 180

QY 181 HFPY 184
 DB 181 HFPY 184

RESULT 14
 CKR5_MOUSE
 ID CKR5_MOUSE STANDARD; PRT; 354 AA.
 AC P51682; Q61867; P97405; O35313; P97308; O35891;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE C-C chemokine receptor type 5 (C-C CKR-5) (CC-CKR-5) (CCR-5) (MIP-1
 alpha receptor).
 DE alpha receptor).
 GN CCR5 OR CKR5.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID:10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-129/SVJ; TISSUE=Spleen;
 RX MEDLINE-96205938; PubMed-6331787;
 RA Boring L., Gosling J., Monteclaro F.S., Lusis A.J., Tsou C.-L.,
 RA Charo I.F.;
 RA "Molecular cloning and functional expression of murine JE (monocyte
 chemoattractant protein 1) and murine macrophage inflammatory protein
 alpha receptors: evidence for two closely linked C-C chemokine
 receptors on chromosome 9.";
 RL J. Biol. Chem. 271:7551-7558(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-C57BL/6 X CHA; TISSUE=Thymus;
 RX MEDLINE-96279910; PubMed-8662890;
 RA Meyer A., Coyne A.J., Proudfoot A.E.I., Wells T.N.C., Power C.A.;
 RA "Cloning and characterization of a novel murine macrophage
 inflammatory protein-1 alpha receptor.";
 RL J. Biol. Chem. 271:14445-14451(1996).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-129/Ola;
 RA Kuziel W.A., Beck M.A., Dawson T.C., Maeda N.;
 RA Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN-C57BL/6, and NIH Swiss; TISSUE=Liver, Kidney, and Spleen;
 RX MEDLINE-98001387; PubMed-9343222;
 RA Kuhlmann S.E., Platt E.J., Kozak S.L., Kabat D.;
 RA "Polymorphisms in the CCR5 genes of African green monkeys and mice
 implicate specific amino acids in infections by simian and human
 immunodeficiency viruses.";
 RL J. Virol. 71:8642-8656(1997).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN-129;
 RX MEDLINE-97404635; PubMed-9261347;
 RA Doranz B.J., Lu Z.H., Rucker J., Zhang T.Y., Sharron M., Cen Y.H.,
 RA Wang Z.X., Guo H.H., Du J.G., Accavitti M.A., Doms R.W., Peiper S.C.;
 RA "Two distinct CCR5 domains can mediate coreceptor usage by human
 immunodeficiency virus type 1.";
 RL J. Virol. 71:6305-6314(1997).
 RN [6]
 RP SEQUENCE FROM N.A.
 RA Guo B., Kuio K., Harada A., Matsushima K.;
 RA Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: RECEPTOR FOR A C-C TYPE CHEMOKINE. BINDS TO MIP-1-ALPHA,
 CC MIP-1-BETA AND RANTES AND SUBSEQUENTLY TRANSDUCES A SIGNAL BY
 CC INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.

CC -1- TISSUE SPECIFICITY: DETECTED IN MONOCYTE/MACROPHAGE CELL LINES,
 CC BUT NOT IN NONHEMATOPOIETIC CELL LINES.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL; U47036; AAC52454.1; -;
 CC DR EMBL; X94151; CRA63867.1; -;
 CC DR EMBL; U86565; AAB37273.1; -;
 CC DR EMBL; U83327; AAC53386.1; -;
 CC DR EMBL; AF022990; AAC53389.1; -;
 CC DR EMBL; AF019772; AAB71183.1; -;
 CC DR EMBL; D83648; BAA12024.1; -;
 CC DR MCD; MGI:107182; Cnkr5.
 CC DR InterPro; IPR000276; GPCR_Rhodopsn.
 CC DR Pfam; PF00001; 7tm_1; 1.
 CC DR PRINTS; PR00237; GPCRHHODPSN.
 CC DR PROSITE; PS00237; G-PROTEIN_RECP_F1_1; 1.
 CC DR PROSITE; PS00262; G-PROTEIN_RECP_F1_2; 1.
 CC KW G-protein coupled receptor; Transmembrane; Glycoprotein; Polymorphism.
 CC FT DOMAIN 1 32
 CC FT TRANSMEM 33 60
 CC FT DOMAIN 61 70
 CC FT CYTOPLASMIC (POTENTIAL).
 CC FT TRANSMEM 71 91
 CC FT DOMAIN 92 104
 CC FT EXTRACELLULAR (POTENTIAL).
 CC FT TRANSMEM 105 126
 CC FT DOMAIN 127 143
 CC FT CYTOPLASMIC (POTENTIAL).
 CC FT TRANSMEM 144 168
 CC FT DOMAIN 169 200
 CC FT EXTRACELLULAR (POTENTIAL).
 CC FT TRANSMEM 201 220
 CC FT DOMAIN 221 237
 CC FT CYTOPLASMIC (POTENTIAL).
 CC FT TRANSMEM 238 262
 CC FT DOMAIN 263 279
 CC FT EXTRACELLULAR (POTENTIAL).
 CC FT TRANSMEM 280 303
 CC FT DOMAIN 304 354
 CC FT CYTOPLASMIC (POTENTIAL).
 CC FT DISULFID 103 180
 CC FT CARBOHYD 270 270
 CC FT VARIANT 11 11 I -> S.
 CC FT VARIANT 62 62 K -> R.
 CC FT VARIANT 66 66 V -> M.
 CC FT VARIANT 97 97 I -> V.
 CC FT VARIANT 109 109 V -> L.
 CC FT VARIANT 156 156 V -> A.
 CC FT VARIANT 160 160 F -> S.
 CC FT VARIANT 185 185 P -> L.
 CC FT VARIANT 213 213 I -> V.
 CC FT VARIANT 318 318 I -> M.
 CC FT VARIANT 337 337 V -> A.
 CC FT CONFLICT 3 3 F -> L (IN REF. 2).
 CC FT CONFLICT 80 80 L -> F (IN REF. 2).
 CC FT CONFLICT 145 145 N -> I (IN REF. 5).
 CC FT CONFLICT 190 190 H -> Y (IN REF. 3).
 CC FT CONFLICT 208 208 P -> S (IN REF. 1).
 CC FT CONFLICT 208 208 P -> S (IN REF. 1).
 CC SQ SEQUENCE 354 AA; 40863 MW; B4A6B942E88F9CF0 CRC64;
 Query Match 70.1%; Score 786; DB 1; Length 354;
 Best Local Similarity 80.1%; Pred. No. 1.le-42;
 Matches 149; Conservative 17; Mismatches 18; Indels 2; Gaps 1;
 QY 1 MDYGVSP--IYDINVTSEPCOKINVQIAARLLPPLSLVIFGFGVGNMVLILLINC 58
 DB 1 MDEGVSPTIYDIYDGNAPCOKINVQIAAQLPPLSLVIFGFGVGNMVLILLISC 60
 QY 59 KRLKSTDIYLLNLTAISDLFLVTFVFAHVAQAQWDFGNMTCOLLTGLYFGFSSIFP 118
 DB 61 KRLKSVTDIYLLNLTAISDLFLVTFVFAHVAQAQWDFGNMTCOLLTGLYFGFSSIFP 120

QY 119 ILLTIDRYLAVHAFALKARTVTRGVTSVITWVAVFASLPGLIIFRSQKGLHYTC 178
 Db 121 ILLTIDRYLAVHAFALKARTVTRGVTSVITWVAVFASLPGLIIFRSQKGLHYTC 180
 QY 179 SSHPY 184
 Db 181 SPHPH 186

RESULT 15
 CKR5_RAT
 ID CKR5_RAT STANDARD; PRT; 354 AA.
 AC 008556;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE C-C chemokine receptor type 5 (C-CR-5) (CCR-5) (MIP-1
 DE alpha receptor).
 GN CCR5 OR CMKR5.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Wistar; TISSUE=Brain;
 RX MEDLINE=98334064; PubMed=9670989;
 RA Spleiss O., Gourmal N., Boddeke H.W.G.M., Sauter A., Fiebich B.L.,
 RA Berger M., Gebicke-Haerter P.J.;
 RT "Cloning of rat HIV-1-chemokine coreceptor CCR5 from microglia and
 RT upregulation of its mRNA in ischemic and endotoxemic rat brain.";
 RL J. Neurosci. Res. 53:16-28(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley;
 RX MEDLINE=98318173; PubMed=9655467;
 RA Jiang Y., Salafranca M.N., Adhikari S., Xia Y., Feng L., Sonntag M.K.,
 RA Defebvre C.M., Pennell N.A., Streit W.J., Harrison J.K.;
 RT "Chemokine receptor expression in cultured glia and rat experimental
 RT allergic encephalomyelitis";
 RL J. Neuroimmunol. 86:1-12(1998).
 CC -1- FUNCTION: RECEPTOR FOR A C-C TYPE CHEMOKINE. BINDS TO MIP-1-ALPHA,
 CC MIP-1-BETA AND RANTES AND SUBSEQUENTLY TRANSDUCES A SIGNAL BY
 CC INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; Y12009; CAAT2737.1; -;
 CC EMBL; U77350; AAC03243.1; -;
 CC InterPro; IPR000276; GPCR_Rhodpsn.
 CC Pfam; PF00001; 7tm_1; 1.
 CC PRINTS; PR00237; GPCRHHODPSN.
 CC PROSITE; PS00237; G-PROTEIN RECEPTOR FL1; 1.
 CC PROSITE; PS00262; G-PROTEIN RECEPTOR FL2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein.
 FT DOMAIN 1 32 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 33 60 1 (POTENTIAL).
 FT DOMAIN 61 70 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 71 91 2 (POTENTIAL).
 FT DOMAIN 92 104 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 105 126 3 (POTENTIAL).
 FT DOMAIN 127 143 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 144 168 4 (POTENTIAL).
 FT DOMAIN 169 200 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 201 220 5 (POTENTIAL).

FT DOMAIN 221 237 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 238 262 6 (POTENTIAL).
 FT DOMAIN 263 279 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 280 303 7 (POTENTIAL).
 FT DOMAIN 304 354 CYTOPLASMIC (POTENTIAL).
 FT DISULFID 103 180 BY SIMILARITY.
 FT CARBOHYD 270 270 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 354 AA; 41030 MW; 77EDB368AA4C868D CRC64;

Query Match 69.1%; Score 775; DB 1; Length 354;
 Best Local Similarity 78.2%; Pred. No. 5.5e-42;
 Matches 147; Conservative 18; Mismatches 21; Indels 2; Gaps 1;

QY 1 MDYQVSSP--TYDINYTSEPCQKINVKQAARLLPPLYSLVFIFGVGNMLVILINC 58
 Db 1 MDPQGSIPTYIIDYDYSMSAPCQKVNKQIAAQLLPPLYSLVFIFGVGNMVFILISC 60
 QY 59 KRKSMTDIYLLNLAISDLFFLLITVPFWAHYAAAQWDFGNMTCOLLGLYFIFGSGIFF 118
 Db 61 KRKSMTDIYLLNLAISDLFFLLITVPFWAHYAAANWFGNIMCKLFTGIYHIGFGGIF 120
 QY 119 ILLTIDRYLAVHAFALKARTVTRGVTSVITWVAVFASLPGLIIFRSQKGLHYTC 178
 Db 121 ILLTIDRYLAVHAFALKARTVTRGVTSVITWVAVFASLPGLIIFRSQKGLHYTC 180
 QY 179 SSHPYIK 186
 Db 181 SPHFLHIQ 188

Search completed: June 3, 2003, 19:22:56
 Job time : 13.2338 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2003, 19:20:46 ; Search time 23.5976 Seconds
(without alignments)
875.891 Million cell updates/sec

Title: US-09-938-703-6
Perfect score: 1122
Sequence: 1 MDQVSPIDYNTSEPC.....AACGHLLGNPKNSASVSK 215

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR73.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Query Score | Match | Length | ID | Description |
|------------|-------------|-------|--------|--------|--------------------|
| 1 | 958 | 85.4 | 352 | A43113 | chemokine (C-C) re |
| 2 | 694 | 61.9 | 360 | JC2443 | chemokine (C-C) re |
| 3 | 694 | 61.9 | 374 | I38450 | chemokine (C-C) re |
| 4 | 598 | 53.3 | 355 | A45177 | chemokine (C-C) re |
| 5 | 573.5 | 51.1 | 355 | I49339 | macrophage inflamm |
| 6 | 547.5 | 48.8 | 360 | A57160 | chemokine (C-C) re |
| 7 | 546.5 | 48.7 | 360 | JC4587 | chemokine (C-C) re |
| 8 | 535.5 | 47.7 | 359 | I49341 | MIP-1 alpha recept |
| 9 | 506.5 | 45.1 | 355 | G02436 | chemokine (C-C) re |
| 10 | 489.5 | 43.6 | 356 | I49340 | MIP-1 alpha recept |
| 11 | 449 | 40.0 | 383 | S55594 | G protein-coupled |
| 12 | 427.5 | 38.1 | 355 | JC5067 | G protein-coupled |
| 13 | 386 | 34.4 | 354 | I58186 | probable G protein |
| 14 | 378 | 33.7 | 344 | JC5942 | chemokine receptor |
| 15 | 365 | 32.5 | 355 | JC4304 | orphan G protein-c |
| 16 | 345.5 | 30.8 | 378 | B55735 | lymphocyte-specifi |
| 17 | 342.5 | 30.5 | 378 | A55735 | G protein-coupled |
| 18 | 333.5 | 29.7 | 378 | A45680 | G protein-coupled |
| 19 | 331.5 | 29.5 | 352 | A45747 | neuropeptide Y/pep |
| 20 | 326 | 29.1 | 353 | S28787 | neuropeptide Y/pep |
| 21 | 325.5 | 29.0 | 352 | G00048 | fusin (LESTRA) - c |
| 22 | 321.5 | 28.7 | 359 | I51372 | angiotensin II rec |
| 23 | 320 | 28.5 | 359 | JC5068 | G protein-coupled |
| 24 | 317 | 28.3 | 369 | JC1194 | angiotensin II rec |
| 25 | 314 | 28.0 | 359 | A43656 | angiotensin II rec |
| 26 | 312 | 27.8 | 359 | JC1104 | angiotensin II rec |
| 27 | 312 | 27.8 | 359 | I39418 | angiotensin II rec |
| 28 | 311 | 27.7 | 359 | JC2134 | angiotensin II rec |
| 29 | 310 | 27.6 | 359 | S15403 | angiotensin II rec |

30 309 27.5 359 2 JH0621 angiotensin II rec
31 308 27.5 359 2 A48857 angiotensin II rec
32 307 27.4 359 2 S44425 angiotensin II rec
33 306 27.3 359 2 JQ1516 angiotensin II rec
34 291 25.9 350 2 JH0621 G protein-coupled
35 287.5 25.6 367 2 JE0349 interferon-inducib
36 283 25.2 362 2 JN0694 angiotensin II rec
37 276 24.6 333 2 I65989 G protein-coupled
38 276 24.6 363 2 I57955 somatostatin recep
39 276 24.6 364 2 JN0763 somatostatin recep
40 274.5 24.5 359 2 A48921 interleukin-8 rece
41 273.5 24.4 384 2 A47249 brain-specific som
42 270.5 24.1 388 2 JN0605 somatostatin recep
43 269 24.0 360 2 A53611 interleukin-8 rece
44 267 23.8 363 2 I57940 somatostatin recep
45 265 23.6 355 2 JQ1231 interleukin-8 rece

ALIGNMENTS

RESULT 1

A43113
chemokine (C-C) receptor 5 - human
N: Alternate names: C-C CKR-5; CCR5
C: Species: Homo sapiens (Man)
C: Date: 12-Jul-1996 #sequence_revision 12-Jul-1996 #text change 20-Jun-2000
C: Accession: A43113; S71808; A58834; A58832; G02653; A58833
R: Samson, M.; Labbe, O.; Mollereau, C.; Vassart, G.; Parmentier, M.
Biochemistry 35, 3362-3367, 1996
A: Title: Molecular cloning and functional expression of a new human CC-chemokine rece
A: Reference number: A43113; MUID:96241590; PMID:8639485
A: Accession: A43113
A: Molecule type: mRNA
A: Residues: 1-352 <SAM>
A: Cross-references: GB:X91492; NID: g1262810; PIDN: CAA62796.1; PID: g1262811
R: Samson, M.; Libert, F.; Doranz, B.J.; Rucker, J.; Liesnard, C.; Farber, C.M.; Sara
M.; Imal, T.; Rana, S.; Yi, Y.; Smyth, R.J.; Collman, R.G.; Doms, R.W.; Vassart, G.;
Nature 382, 722-725, 1996
A: Title: Resistance to HIV-1 infection in caucasian individuals bearing mutant allele
A: Reference number: S71808; MUID:96345670; PMID:8751444
A: Accession: S71808
A: Status: nucleic acid sequence not shown; not compared with conceptual translation
A: Molecule type: DNA
A: Residues: 182-206; 207-230 <SAM2>
A: Accession: A58834
A: Status: nucleic acid sequence not shown; not compared with conceptual translation
A: Molecule type: DNA
A: Residues: 1-184; 'KDSHLGAGPAACHGHLILGNPKNSASVSK' <SAM3>
A: Cross-references: GB:X93393; NID: g1524062; PIDN: CAA67767.1; PID: g1524063
A: Note: this frameshift mutation results in a non-functional receptor but confers a
nd may have had a selective advantage by conferring resistance to Yersinia plague in
R: Combadiere, C.; Ahuja, S.K.; Tiffany, H.L.; Murphy, P.M.
J. Leukoc. Biol. 60, 147-152, 1996
A: Title: Cloning and functional expression of CC CKR5, a human monocyte CC chemokine
A: Reference number: A58832; MUID:96295970; PMID:8699119
A: Accession: A58832
A: Molecule type: mRNA
A: Residues: 1-352 <COM1>
A: Cross-references: GB:U057840; NID: g1502408; PIDN: AAB17071.1; PID: g1502409
A: Experimental source: clone 8, endotoxin-stimulated peripheral blood monocytes
R: Combadiere, C.
submitted to the EMBL Data Library, May 1996
A: Reference number: H01541
A: Accession: G02653
A: Status: translated from GB/EMBL/DBJ
A: Molecule type: mRNA
A: Residues: 1-89; 'L' / 91-352 <COM2>
A: Cross-references: EMBL:U57840
R: Raport, C.J.; Gosling, J.; Schweickart, V.L.; Gray, P.W.; Charo, I.F.
J. Biol. Chem. 271, 17161-17166, 1996
A: Title: Molecular cloning and functional characterization of a novel human CC chemok
A: Reference number: A58833; MUID:96291862; PMID:8663314

A:Accession: A58833
 A:Molecule type: mRNA
 A:Residues: 1-352 <RAP>
 A:Cross-references: GB:U54994; NID:g1457945; PIDN:AAC50598.1; PID:g1457946
 C:Comment: This is a receptor for chemokines MIP-1alpha (see PIR:A30574), MIP-1beta (see PIR:A30575) and dual-tropic strains of HIV-1 bind to a complex of chemokine receptors.
 C:Genetics:
 A:Gene: GDB:CMKBR5; CCR5; CCR-5; CC-CR-5; CCR5; ChemR13
 A:Cross-references: GDB:1230510; OMIM:601373
 A:Map position: 3p21-3p21
 C:Function:
 A:Description: G protein-coupled receptor for chemokines MIP-1alpha, MIP-1beta and RANTES.
 A:Note: probably acts to control granulocyte proliferation and differentiation
 C:Superfamily: vertebrate rhodopsin
 C:Keywords: AIDS; G protein-coupled receptor; glycoprotein; phosphoprotein; transmembrane protein
 F:32-56/Domain: transmembrane #status predicted <TM1>
 F:67-87/Domain: transmembrane #status predicted <TM2>
 F:103-124/Domain: transmembrane #status predicted <TM3>
 F:142-166/Domain: transmembrane #status predicted <TM4>
 F:193-218/Domain: transmembrane #status predicted <TM5>
 F:236-257/Domain: transmembrane #status predicted <TM6>
 F:285-300/Domain: transmembrane #status predicted <TM7>
 F:20-269,101-178/Disulfide bonds: #status predicted
 F:268/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:336,337,342/Binding site: phosphate (Ser) (covalent) #status predicted
 F:340,343/Binding site: phosphate (Thr) (covalent) #status predicted

Query Match 85.4%; Score 958; DB 2; Length 352;
 Best Local Similarity 100.0%; Pred. No. 1.2e-80;
 Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDYQSSPIYDINYYSEPCQKINVKQIAARLLPPLYSILVIFGFGNMVLVILLINCKR 60
 (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||)
 Db 1 MDYQSSPIYDINYYSEPCQKINVKQIAARLLPPLYSILVIFGFGNMVLVILLINCKR 60
 (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||)

QY 61 LKSMTDIYLLNLALSDLFLLVPPFWAHYAAQWDFGNTMCQLTGLYFGFGFIIFII 120
 (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||)
 Db 61 LKSMTDIYLLNLALSDLFLLVPPFWAHYAAQWDFGNTMCQLTGLYFGFGFIIFII 120
 (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||)

QY 121 LFTIDRYLAVHAFALKARTVGVTVSVITWVAVFASLPGLIIFRSQKGLHYTCSS 180
 (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||)
 Db 121 LFTIDRYLAVHAFALKARTVGVTVSVITWVAVFASLPGLIIFRSQKGLHYTCSS 180
 (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||)

QY 181 HPPY 184
 (|||||)
 Db 181 HPPY 184
 (|||||)

RESULT 2
 JC2443
 chemokine (C-C) receptor 2, splice form B - human
 N:Alternate names: C-C CR-2; monocyte chemoattractant protein 1 receptor; monocyte chemokine receptor
 C:Species: Homo sapiens (man)
 C:Date: 21-Feb-1995 #sequence_revision 05-Apr-1995 #text_change 20-Jun-2000
 C:Accession: JC2443; J38463
 R:Yamagami, S.; Tokuda, Y.; Ishii, K.; Tanaka, H.; Endo, N.
 Biochem. Biophys. Res. Commun. 202, 1156-1162, 1994
 A:Title: cDNA cloning and functional expression of a human monocyte chemoattractant protein 1 receptor
 A:Reference number: JC2443; MUID:94324942; PMID:8048929
 A:Accession: JC2443
 A:Molecule type: mRNA
 A:Residues: 1-360 <YAM>
 A:Cross-references: DDBJ:D29984; NID:g531246; PIDN:BAA06253.1; PID:g531247
 R:Charo, I.F.; Myers, S.J.; Herman, A.; Franci, C.; Connolly, A.J.; Coughlin, S.R.
 Proc. Natl. Acad. Sci. U.S.A. 91, 2752-2756, 1994
 A:Title: Molecular cloning and functional expression of two monocyte chemoattractant proteins 1 and 2
 A:Reference number: A53477; MUID:94195821; PMID:8146186
 A:Accession: J38463
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-360 <RES>
 A:Cross-references: EMBL:U03905; NID:g472557; PIDN:AAA19120.1; PID:g472558
 C:Genetics:

A:Gene: GDB:CMKBR2
 A:Cross-references: GDB:337364; OMIM:601267
 A:Map position: 3p21-3p21
 C:Superfamily: vertebrate rhodopsin
 C:Keywords: alternative splicing; G protein-coupled receptor; glycoprotein; transmembrane protein
 F:43-70/Domain: transmembrane #status predicted <TM1>
 F:81-100/Domain: transmembrane #status predicted <TM2>
 F:115-136/Domain: transmembrane #status predicted <TM3>
 F:154-178/Domain: transmembrane #status predicted <TM4>
 F:207-226/Domain: transmembrane #status predicted <TM5>
 F:244-268/Domain: transmembrane #status predicted <TM6>
 F:287-309/Domain: transmembrane #status predicted <TM7>
 F:14/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:113-190/Disulfide bonds: #status predicted

Query Match 61.9%; Score 694; DB 2; Length 360;
 Best Local Similarity 75.9%; Pred. No. 2.4e-56;
 Matches 132; Conservative 17; Mismatches 23; Indels 2; Gaps 1;

QY 10 YDINYYTSEPCQKINVKQIAARLLPPLYSILVIFGFGNMVLVILLINCKRLKSMTDIYL 69
 (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||)
 Db 24 FDYDYGAPCHKFDVKQIGQALLPPLYSILVIFGFGNMVLVILLINCKRLKCLDIYL 81
 (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||)

QY 70 LNLALSDLFLLVPPFWAHYAAQWDFGNTMCQLTGLYFGFGFIIFIIITIDRYLA 129
 (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||)
 Db 82 LNLALSDLFLLVPPFWAHYAAQWDFGNTMCQLTGLYFGFGFIIFIIITIDRYLA 141
 (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||)

QY 130 VYHAFALKARTVGVTVSVITWVAVFASLPGLIIFRSQKGLHYTCSSHPF 183
 (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||)
 Db 142 IVHAFALKARTVGVTVSVITWVAVFASLPGLIIFRSQKGLHYTCSSHPF 195
 (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||)

RESULT 3
 I38450
 chemokine (C-C) receptor 2, splice form A - human
 N:Alternate names: C-C CR-2; monocyte chemoattractant protein 1 receptor; monocyte chemokine receptor
 C:Species: Homo sapiens (man)
 C:Date: 16-Feb-1995 #sequence_revision 16-Feb-1996 #text_change 13-Aug-1999
 C:Accession: J38450
 R:Charo, I.F.; Myers, S.J.; Herman, A.; Franci, C.; Connolly, A.J.; Coughlin, S.R.
 Proc. Natl. Acad. Sci. U.S.A. 91, 2752-2756, 1994
 A:Title: Molecular cloning and functional expression of two monocyte chemoattractant proteins 1 and 2
 A:Reference number: A53477; MUID:94195821; PMID:8146186
 A:Accession: J38450
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-374 <RES>
 A:Cross-references: EMBL:U03882; NID:g472555; PIDN:AAA19119.1; PID:g472556
 C:Genetics:

A:Gene: GDB:CMKBR2
 A:Cross-references: GDB:337364; OMIM:601267
 A:Map position: 3p21-3p21
 C:Superfamily: vertebrate rhodopsin
 C:Keywords: alternative splicing; G protein-coupled receptor; glycoprotein; transmembrane protein
 F:44-68/Domain: transmembrane #status predicted <TM1>
 F:79-99/Domain: transmembrane #status predicted <TM2>
 F:115-136/Domain: transmembrane #status predicted <TM3>
 F:154-178/Domain: transmembrane #status predicted <TM4>
 F:208-226/Domain: transmembrane #status predicted <TM5>
 F:244-265/Domain: transmembrane #status predicted <TM6>
 F:292-309/Domain: transmembrane #status predicted <TM7>
 F:14/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:32-277,113-190/Disulfide bonds: #status predicted

Query Match 61.9%; Score 694; DB 2; Length 374;
 Best Local Similarity 75.9%; Pred. No. 2.5e-56;
 Matches 132; Conservative 17; Mismatches 23; Indels 2; Gaps 1;

QY 10 YDINYYTSEPCQKINVKQIAARLLPPLYSILVIFGFGNMVLVILLINCKRLKSMTDIYL 69
 (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||)
 Db 24 FDYDYGAPCHKFDVKQIGQALLPPLYSILVIFGFGNMVLVILLINCKRLKCLDIYL 81
 (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||)

QY 70 LNLALSDLFLLVPPFWAHYAAQWDFGNTMCQLTGLYFGFGFIIFIIITIDRYLA 129
 (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||)

Db
|||: 186 HEPH 189

C;superfamily: vertebrata

DB L35 IVHAFALRRARTVTLGLITTSITWALAILASMPALYFFKQWEEFTHRT'CSPPFPY 189

C;Genetics:

Quarry Match

CONCLUSIONS

A;Cross-references: EMBL:D8406; NID:g981551; PID:g981552
C:Superfamily: vertebrate rhodopsin

Query Match 47.7%; Score 535.5; DB 2; Length 359;
Best Local Similarity 57.9%; Pred. No. 9e-42;
Matches 99; Conservative 31; Mismatches 40; Indels 1; Gaps 1;

QY 14 YTSPECKINVKQAARLLPPYSLVFTFGVGNMLVLILNCRKLSMTDIYLLNLIA 73
 | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 22 YEAWPCEKVRIKELGWSLLDPPLSYLVFTIIGLGNNMVLIIKYRKLIQIMTIYFNLA 81

QY 74 ISDLFFLTVPFWAHYA-AOWDGMTWCQLLTGLYFLGFSGFIETILLTDRYLANVH 132
 ||||| ||||| ||||| ||||| :| |||| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 82 ISDLFFLTVPFVIHVHWLNWNGHGVMCKLUSGYIALYSEIFFILLTDRYLAIIVH 141

QY 133 AVFKARTVTGTGVSTVTVVAASLPGLITFRSQEGLHYTCSSHP 183
 ||||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 142 AVFAARATVTETAITSIITWGTLAALPEFIPESQDSFGFSCSRYP 192

RESULT 9
GO2436
chemokine (C-C) receptor 3 - human
N:Alternate names: C-C CKR-3
C:Species: Homo sapiens (man)
C:Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 04-Mar-2000
C:Accession: GO2436; A57237
R:Ponath, P.D.
Submitted to the EMBL Data Library, February 1996
A:Reference number: H01272
A:Accession: GO2436
A>Status: translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-355 <PON>
A:Cross-references: EMBL:U49727; NID:g1477560; PID:NAB09726.1; PID:g1477561
R:Combadiere, C.; Ahuja, S.K.; Murphy, P.M.
J. Biol. Chem. 270, 16491-16494, 1995
A:Title: Cloning and functional expression of a human eosinophil CC chemokine receptor
A:Reference number: A57237; MUID:95348056; PMID:7622448
A:Accession: A57237
A>Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-106,'N',108-275','S',277-280,'R',282-355 <COM>
A:Cross-references: GB:U28694; NID:g1199579; PID:NAC50469.1; PID:g1199580
A>Note: The translated sequence in GenBank entry HSU28694, release 113.0, PID:NAC50469.1
C:Genetics:
A:Gene: GDB:CMKBR3
A:Cross-references: GDB:579624; OMIM:601268
A:Map position: 3p21-3p21
C:Superfamily: vertebrate rhodopsin
C:Keywords: G protein-coupled receptor; glycoprotein; phosphoprotein; transmembrane p
F:36-60/Domain: transmembrane #status predicted <TM1>
F:71-91/Domain: transmembrane #status predicted <TM2>
F:108-129/Domain: transmembrane #status predicted <TM3>
F:147-171/Domain: transmembrane #status predicted <TM4>
F:205-223/Domain: transmembrane #status predicted <TM5>
F:240-261/Domain: transmembrane #status predicted <TM6>
F:288-305/Domain: transmembrane #status predicted <TM7>
F:247-106-183/Disulfide bonds: #status predicted
F:345/Binding site: phosphate (Ser) (covalent) (by casein kinase II) #status predicted

Query Match 45.1%; Score 506.5; DB 2; Length 355;
Best Local Similarity 57.6%; Pred. No. 4.2e-39;
Matches 95; Conservative 32; Mismatches 37; Indels 1; Gaps 1;

QY 20 CORINKQTAAARLLPPYSLVFTFGVGNMLVLILNCRKLSMTDIYLLNLIAISDLFF 79
 | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 24 CERADRALMAQFPPLPSYLVFTVGLLGNVVMIKYRLRMNIYLLNLIAISDLLF 83

QY 80 LLVTPFWAHYAAA-OWDGNTWCQLLTGLYFLGFSGFIETILLTDRYLANVHAIFALK 138
 ||||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 84 LVLPFWIHVRGNWVFHGKCLKSFGFYHTGLSYSEIFFILLTDRYLAIVHANPAFLR 143

Search completed: June 3, 2003, 19:24:27
Job time : 24.5976 secs

RESULT 15

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2003, 19:17:26 ; Search time 7.93902 Seconds
(without alignments)
520.312 Million cell updates/sec

Title: US-09-938-703-6_COPY_185_215
Perfect score: 164
Sequence: 1 IKDHLGAGPAAACHGILLGNPKNSASVSK 31

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues
Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : A.Geneseq.101002.*
- 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
 - 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
 - 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
 - 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
 - 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
 - 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
 - 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
 - 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
 - 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
 - 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
 - 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
 - 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
 - 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
 - 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
 - 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
 - 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
 - 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
 - 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
 - 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
 - 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
 - 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
 - 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
 - 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------|
| 1 | 164 | 100.0 | 215 | 18 | AAW27408 |
| 2 | 164 | 100.0 | 215 | 20 | AAW88238 |
| 3 | 55 | 33.5 | 84 | 22 | ABB36998 |
| 4 | 55 | 33.5 | 84 | 22 | AAW57726 |
| 5 | 55 | 33.5 | 84 | 22 | AAW70138 |
| 6 | 55 | 33.5 | 84 | 23 | ABG39777 |
| 7 | 55 | 33.5 | 94 | 22 | ABG23873 |
| 8 | 54 | 32.9 | 276 | 22 | ABB69618 |
| 9 | 53.5 | 32.6 | 189 | 22 | ABG27770 |
| 10 | 53.5 | 32.6 | 197 | 22 | ABG27246 |

| | | | | | | |
|----|------|------|------|----|----------|--------------------|
| 11 | 52.5 | 32.0 | 216 | 21 | AAG18166 | Arabidopsis thalia |
| 12 | 52.5 | 32.0 | 265 | 21 | AAG18165 | Arabidopsis thalia |
| 13 | 52.5 | 32.0 | 270 | 21 | AAG28063 | Arabidopsis thalia |
| 14 | 52.5 | 32.0 | 478 | 21 | AAG28062 | Arabidopsis thalia |
| 15 | 51 | 31.1 | 193 | 22 | ABG21033 | Novel human diagno |
| 16 | 49 | 29.9 | 278 | 21 | AAG22079 | Arabidopsis thalia |
| 17 | 49 | 29.9 | 278 | 21 | AAG43094 | Arabidopsis thalia |
| 18 | 49 | 29.9 | 290 | 21 | AAG22078 | Arabidopsis thalia |
| 19 | 49 | 29.9 | 290 | 21 | AAG43093 | Arabidopsis thalia |
| 20 | 49 | 29.9 | 331 | 21 | AAG22077 | Arabidopsis thalia |
| 21 | 49 | 29.9 | 331 | 21 | AAG43092 | Arabidopsis thalia |
| 22 | 48.5 | 29.6 | 698 | 22 | ABB59005 | Drosophila melanog |
| 23 | 48 | 29.3 | 101 | 22 | AAO07177 | Human polypeptide |
| 24 | 48 | 29.3 | 186 | 23 | ABP42972 | Human ovarian anti |
| 25 | 48 | 29.3 | 297 | 23 | AAU77123 | Human protease #2. |
| 26 | 48 | 29.3 | 320 | 22 | AAU19238 | Human G protein-co |
| 27 | 48 | 29.3 | 340 | 22 | AAU40849 | Propionibacterium |
| 28 | 48 | 29.3 | 451 | 23 | AAU77122 | Human protease #1. |
| 29 | 48 | 29.3 | 486 | 23 | AAU77124 | Human protease #3. |
| 30 | 48 | 29.3 | 828 | 22 | ABB63141 | Drosophila melanog |
| 31 | 48 | 29.3 | 1186 | 23 | AAU72895 | Human metalloprote |
| 32 | 48 | 29.3 | 1189 | 23 | AAU74750 | Human protease PRT |
| 33 | 48 | 29.3 | 1210 | 23 | AAU85415 | Human protein NOV1 |
| 34 | 48 | 29.3 | 1213 | 23 | AAU77128 | Human protease #7. |
| 35 | 48 | 29.3 | 1216 | 23 | AAU77127 | Human protease #6. |
| 36 | 48 | 29.3 | 1219 | 23 | AAU77126 | Human protease #5. |
| 37 | 48 | 29.3 | 1222 | 23 | AAU77125 | Human metalloprote |
| 38 | 48 | 29.3 | 1223 | 23 | ABG30863 | Human metalloprote |
| 39 | 48 | 29.3 | 1223 | 23 | AAU79747 | Human ADAM-TS 1 pr |
| 40 | 48 | 29.3 | 1232 | 23 | AAU77130 | Human protease #9. |
| 41 | 48 | 29.3 | 1235 | 23 | AAU77129 | Human protease #8. |
| 42 | 48 | 29.3 | 1249 | 23 | AAU77132 | Human protease #11 |
| 43 | 48 | 29.3 | 1252 | 23 | AAU77131 | Human protease #10 |
| 44 | 47.5 | 29.0 | 146 | 22 | ABG19311 | Novel human diagno |
| 45 | 47.5 | 29.0 | 971 | 22 | ABB63091 | Drosophila melanog |

ALIGNMENTS

RESULT 1
AAW27408
ID AAW27408 standard; Protein; 215 AA.
XX AC AAW27408;
XX DT 14-APR-1998 (first entry)
XX DE Inactive human CCR5.
XX KW Inactive; human Cys-Cys chemokine receptor 5; CCR5;
KW human immunodeficiency virus; type 1; type 2; HIV-1; HIV-2;
KW predisposition; resistance; diagnosis; treatment; prevention;
KW inflammatory disease; rheumatoid arthritis; glomerulonephritis;
KW asthma; idiopathic pulmonary fibrosis; psoriasis; viral infection;
KW cancer; atherosclerosis; autoimmune disorder.
XX OS Homo sapiens.
XX PN WO9732019-A2.
XX PD 04-SEP-1997.
XX PF 28-FEB-1997; 97WO-BE00023.
XX PR 06-AUG-1996; 96EP-0870102.
XX PR 01-MAR-1996; 96EP-0870021.
XX PA (EURO-) EUROSCREEN SA.
XX PI Libert F, Parmentier M, Samson M, Vassart G;
XX WPI; 1997-479829/44.

XX N-PSDB; AAT90118.
 PT Active and inactive forms of human CC chemokine receptor CCR-5 -
 PT useful to diagnose, prevent and/or treat inflammatory disorders,
 PT autoimmune disease and viral infection
 XX
 PS Claim 7; Fig 1d-e; 94pp; English.
 XX
 CC The present sequence is an inactive human CC (Cys-Cys)
 CC chemokine receptor 5 (CCR5), which lacks the last 3 transmembrane
 CC regions and the regions involved in G protein-coupling. CCR5 or
 CC its cDNA can be used to diagnose, treat and/or prevent inflammatory
 CC diseases, e.g. rheumatoid arthritis, glomerulonephritis, asthma,
 CC idiopathic pulmonary fibrosis and psoriasis, viral infections,
 CC especially human immunodeficiency virus type 1 or type 2 (HIV-1 or
 CC HIV-2) infection, cancer, atherosclerosis and autoimmune disorders.
 CC Subjects that express the inactive receptor have a predisposition,
 CC or resistance to HIV-1 and/or HIV-2.
 XX
 SQ Sequence 215 AA;
 Query Match 100.0%; Score 164; DB 18; Length 215;
 Best Local Similarity 100.0%; Pred. No. 4.7e-15;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 IKDSLHGAGPAAACHGHLILGNPKNSASVSK 31
 ||||||||||||||||||||||||||||||||
 Db 185 IKDSLHGAGPAAACHGHLILGNPKNSASVSK 215
 RESULT 2
 AAW8238
 ID AAW8238 standard; Protein; 215 AA.
 XX
 AC AAW8238;
 XX
 DT 15-MAR-1999 (first entry)
 XX
 DE HIV-1 co-receptor CCR5 variant CCR5-delta32.
 XX
 KW HIV-1; CCR5; CCR5-delta32; co-receptor; infection; diagnosis; AIDS;
 KW gene therapy; human.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Domain 32..56
 FT /note= "transmembrane domain 1"
 FT Domain 67..87
 FT /note= "transmembrane domain 2"
 FT Domain 103..124
 FT /note= "transmembrane domain 3"
 FT Domain 142..167
 FT /note= "transmembrane domain 4"
 XX
 PN WO9854317-A1.
 XX
 PD 03-DEC-1998.
 XX
 PF 29-MAY-1998; 98WO-EP03437.
 XX
 PR 30-MAY-1997; 97US-0048057.
 XX
 PA (MOND-) FOND MONDIALE RECH & PREVENTION SIDA.
 XX
 PI Arenzana Siesdedos F, Beretta A, Braun J, Quillent C;
 XX
 DR WPI; 1999-059835/05.
 DR N-PSDB; AAW84159.
 XX
 PT New CCR5 variant protein of the HIV-1 co-receptor - useful in
 PT developing resistance of CCR5-expressing cells to HIV-1 infection
 XX

PS Disclosure; Page 38-39; 55pp; English.
 XX
 CC This is the amino acid sequence of a CCR5 variant protein,
 CC designated CCR5-delta32, that includes the first 4 transmembrane
 CC domains of wild-type CCR5 (see AAW88232), but lacks transmembrane
 CC domains 5-7. CCR5 serves as a co-receptor for infection by
 CC macrophage-tropic (M-tropic) strains of HIV-1. Individuals
 CC homozygous for the CCR5-delta32 mutation are resistant to HIV-1
 CC infection, but heterozygous individuals are susceptible. The
 CC invention additionally relates to the identification of variant
 CC CCR5s (see AAW88231), which lacks transmembrane domains 3-7 of
 CC CCR5. The detection of CCR5 variants may be used to identify
 CC individuals at lower risk of infection relative to the general
 CC population who, if infected, may exhibit slower progression to
 CC AIDS. Probes and primers (see AAW84127-36) are provided for use in
 CC diagnostic methods for detecting the presence of such variants. A
 CC method is provided for inhibiting HIV-1 infection of a cell
 CC expressing the CCR5 receptor. This involves introducing a nucleic
 CC acid encoding a CCR5 variant into the cell, thereby reducing the
 CC number of functional CCR5 molecules present on the cell surface.
 XX
 SQ Sequence 215 AA;
 Query Match 100.0%; Score 164; DB 20; Length 215;
 Best Local Similarity 100.0%; Pred. No. 4.7e-15;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 IKDSLHGAGPAAACHGHLILGNPKNSASVSK 31
 ||||||||||||||||||||||||||||||||
 Db 185 IKDSLHGAGPAAACHGHLILGNPKNSASVSK 215
 RESULT 3
 ABB36998
 ID ABB36998 standard; Peptide; 84 AA.
 XX
 AC ABB36998;
 XX
 DT 04-FEB-2002 (first entry)
 XX
 DE Peptide #4504 encoded by human foetal liver single exon probe.
 XX
 KW Human; foetal liver; gene expression; single exon nucleic acid probe.
 XX
 OS Homo sapiens.
 XX
 PN WO200157277-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00669.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 FI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-483447/52.
 XX
 PT Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human fetal liver -
 XX
 FS Claim 27; SEQ ID NO 29633; 639pp + sequence listing; English.
 XX
 CC The invention relates to a single exon nucleic acid probe for
 CC measuring human gene expression in a sample derived from human foetal

CC liver. The single exon nucleic acid probes may be used for predicting,
 CC measuring and displaying gene expression in samples derived from human
 CC fetal liver. The present sequence is a peptide encoded by a single exon
 CC nucleic acid probe of the invention.

CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at http://wipo.int/pub/published_pct_sequences.

XX SQ Sequence 84 AA;

Query Match 33.5%; Score 55; DB 22; Length 84;

Best Local Similarity 50.0%; Pred. No. 3.3;

Matches 12; Conservative 1; Mismatches 5; Indels 6; Gaps 1;

QY 3 DSHLGAGPAAACHGHLILGNPKNS 26

Db 48 DSHLGAGPAAACHGHLILGNPKNS 65

RESULT 4

AAM57726

ID AAM57726 standard; Protein; 84 AA.

XX AC AAM57726;

XX DT 05-NOV-2001 (first entry)

XX DE Human brain expressed single exon probe encoded protein SEQ ID NO: 29831.

XX KW Human; brain expressed exon; gene expression analysis; probe;

XX KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;

XX KW epilepsy; cancer.

XX OS Homo sapiens.

XX PN WO200157275-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US00667.

XX PR 04-FEB-2000; 2000US-0180312.

XX PR 26-MAY-2000; 2000US-0207456.

XX PR 30-JUN-2000; 2000US-0608408.

XX PR 03-AUG-2000; 2000US-0632366.

XX PR 21-SEP-2000; 2000US-0234687.

XX PR 27-SEP-2000; 2000US-0236359.

XX PR 04-OCT-2000; 2000GB-0024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX PD WPI; 2001-483446/52.

XX PF Single exon nucleic acid probes for analyzing gene expression in human

XX PT brains -

XX PS Example 4; SEQ ID NO: 29831; 650pp + Sequence Listing; English.

XX CC The present invention provides a number of single exon nucleic acid

XX CC probes which are derived from genomic sequences expressed in the human

XX CC brain. They can be used to measure gene expression in brain cell samples,

XX CC which may enable the diagnosis and improved treatment of nervous system

XX CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,

XX CC epilepsy and cancers. The present sequence is a protein encoded by one of

XX CC the probes of the invention.

XX CC Sequence 84 AA;

XX Query Match 33.5%; Score 55; DB 22; Length 84;

XX Best Local Similarity 50.0%; Pred. No. 3.3;

XX Matches 12; Conservative 1; Mismatches 5; Indels 6; Gaps 1;

QY 3 DSHLGAGPAAACHGHLILGNPKNS 26

Db 48 DSHLGAGPAAACHGHLILGNPKNS 65

RESULT 5

AAM70138

ID AAM70138 standard; Protein; 84 AA.

XX AC AAM70138;

XX DT 06-NOV-2001 (first entry)

XX DE Human bone marrow expressed probe encoded protein SEQ ID NO: 30444.

XX KW Human; bone marrow expressed exon; gene expression analysis; probe;

XX KW microarray; cancer; leukaemia; lymphoma; myeloma.

XX OS Homo sapiens.

XX PN WO200157276-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US00668.

XX PR 04-FEB-2000; 2000US-0180312.

XX PR 26-MAY-2000; 2000US-0207456.

XX PR 30-JUN-2000; 2000US-0608408.

XX PR 03-AUG-2000; 2000US-0632366.

XX PR 21-SEP-2000; 2000US-0234687.

XX PR 27-SEP-2000; 2000US-0236359.

XX PR 04-OCT-2000; 2000GB-0024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX PD WPI; 2001-488900/53.

XX PF Human genome-derived single exon nucleic acid probes useful for

XX PT analyzing gene expression in human bone marrow -

XX PS Example 4; SEQ ID NO: 30444; 658pp + Sequence Listing; English.

XX CC The present invention provides a number of single exon nucleic acid

XX CC probes which are derived from genomic sequences expressed in the human

XX CC bone marrow. They can be used to measure gene expression in bone marrow

XX CC samples, which may enable the improved diagnosis and treatment of cancers

XX CC such as lymphoma, leukaemia and myeloma. The present sequence is a

XX CC protein encoded by one of the probes of the invention.

XX CC Sequence 84 AA;

XX Query Match 33.5%; Score 55; DB 22; Length 84;

XX Best Local Similarity 50.0%; Pred. No. 3.3;

XX Matches 12; Conservative 1; Mismatches 5; Indels 6; Gaps 1;

QY 3 DSHLGAGPAAACHGHLILGNPKNS 26

Db 48 DSHLGAGPAAACHGHLILGNPKNS 65

RESULT 6

ABG39777

ID ABG39777 standard; Peptide; 84 AA.

XX AC ABG39777;

XX DT 19-AUG-2002 (first entry)

XX DE Human peptide encoded by genome-derived single exon probe SEQ ID 29442.

RESULT 10

ABG27246
ID ABG27246 standard; Protein; 197 AA.XX AC ABG27246;
XX DT 18-FEB-2002 (first entry)XX DE Novel human diagnostic protein #27237.
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KW food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US08631.

XX PR 31-MAR-2000; 2000US-0540217.

XX PR 23-AUG-2000; 2000US-0649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX DR N-PSDB; AAS91433.

XX DR WPI; 2001-639362/73.

XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity -
XX PS Claim 20; SEQ ID No 57605; 103pp; English.XX CC The invention relates to isolated polynucleotide (I) and
XX CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX CC and gene mapping, and in recombinant production of (II). The
XX CC polynucleotides are also used in diagnostics as expressed sequence tags
XX CC for identifying expressed genes. (I) is useful in gene therapy techniques
XX CC to restore normal activity of (II) or to treat disease states involving
XX CC (II). (II) is useful for generating antibodies against it, detecting or
XX CC quantitating a polypeptide in tissue, as molecular weight markers and as
XX CC a food supplement. (II) and its binding partners are useful in medical
XX CC imaging of sites expressing (II). (I) and (II) are useful for treating
XX CC disorders involving aberrant protein expression or biological activity.
XX CC The polypeptide and polynucleotide sequences have applications in
XX CC diagnostics, forensics, gene mapping, identification of mutations
XX CC responsible for genetic disorders or other traits to assess biodiversity
XX CC and to produce other types of data and products dependent on DNA and
XX CC amino acid sequences. ABG00010-ABG30377 represent novel human
XX CC diagnostic amino acid sequences of the invention.
XX CC Note: The sequence data for this patent did not appear in the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pot_sequences.

XX SQ Sequence 197 AA;

Query Match 32.6%; Score 53.5; DB 22; Length 197;

Best Local Similarity 43.8%; Pred. No. 14;
Matches 14; Conservative 5; Mismatches 10; Indels 3; Gaps 2;

QY 2 KDSHLAGAPAAACHGHLILGNP--KNSASVSK 31

Db 132 RDS-VGVGPSPCAGYNLLINPPINKNESAKE 162

RESULT 11

AAG18166

ID AAG18166 standard; Protein; 216 AA.

XX AC AAG18166;
XX DT 17-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 19466.

XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence.

XX OS Arabidopsis thaliana.

XX PN EPI033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX PR 25-FEB-1999; 99US-0121025.

XX PR 05-MAR-1999; 99US-0123180.

XX PR 09-MAR-1999; 99US-0123548.

XX PR 23-MAR-1999; 99US-0125788.

XX PR 25-MAR-1999; 99US-0126264.

XX PR 29-MAR-1999; 99US-0126785.

XX PR 01-APR-1999; 99US-0127462.

XX PR 06-APR-1999; 99US-0128234.

XX PR 08-APR-1999; 99US-0128714.

XX PR 16-APR-1999; 99US-0129845.

XX PR 19-APR-1999; 99US-0130077.

XX PR 21-APR-1999; 99US-0130449.

XX PR 23-APR-1999; 99US-0130510.

XX PR 23-APR-1999; 99US-0130891.

XX PR 28-APR-1999; 99US-0131449.

XX PR 30-APR-1999; 99US-0132048.

XX PR 30-APR-1999; 99US-0132407.

XX PR 04-MAY-1999; 99US-0132484.

XX PR 05-MAY-1999; 99US-0132485.

XX PR 06-MAY-1999; 99US-0132486.

XX PR 07-MAY-1999; 99US-0132487.

XX PR 11-MAY-1999; 99US-0132863.

XX PR 14-MAY-1999; 99US-0134218.

XX PR 14-MAY-1999; 99US-0134219.

XX PR 14-MAY-1999; 99US-0134221.

XX PR 14-MAY-1999; 99US-0134370.

XX PR 18-MAY-1999; 99US-0134768.

XX PR 19-MAY-1999; 99US-0134941.

XX PR 20-MAY-1999; 99US-0135124.

XX PR 21-MAY-1999; 99US-0135353.

XX PR 24-MAY-1999; 99US-0135629.

XX PR 25-MAY-1999; 99US-0136021.

XX PR 27-MAY-1999; 99US-0136392.

XX PR 28-MAY-1999; 99US-0136782.

XX PR 01-JUN-1999; 99US-0137222.

XX PR 03-JUN-1999; 99US-0137528.

XX PR 04-JUN-1999; 99US-0137502.

XX PR 07-JUN-1999; 99US-0137724.

XX PR 08-JUN-1999; 99US-0138094.

XX PR 10-JUN-1999; 99US-0138540.

XX PR 10-JUN-1999; 99US-0138847.

XX PR 14-JUN-1999; 99US-0139119.

XX PR 16-JUN-1999; 99US-0139452.

XX PR 16-JUN-1999; 99US-0139453.

XX PR 17-JUN-1999; 99US-0139492.

XX PR 18-JUN-1999; 99US-0139454.

XX PR 18-JUN-1999; 99US-0139455.

XX PR 18-JUN-1999; 99US-0139456.

XX PR 18-JUN-1999; 99US-0139457.

XX PR 18-JUN-1999; 99US-0139458.

XX PR 18-JUN-1999; 99US-0139459.

XX PR 18-JUN-1999; 99US-0139460.

PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145313.
PR 27-JUL-1999; 99US-0145318.
PR 27-JUL-1999; 99US-0145919.
PR 27-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.

PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 32.0%; Score 52.5; DB 21; Length 216;
Best Local Similarity 45.5%; Pred. No. 22;
Matches 10; Conservative 4; Mismatches 7; Indels 1; Gaps 1;

QY 2 KDSHLGAGPAACHGILLGNP 23
::: ||| |||||
Db 84 RNDNGYGGPGG-HGHLMAGNP 104

RESULT 12
AAG18165
ID AAG18165 standard; Protein: 265 AA.

XX AAG18165;

AC AAG18165;
XX 17-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 19465.
DE

XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
OS
PN EPI033405-A2.
XX
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.

PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 32.0%; Score 52.5; DB 21; Length 265;
Best Local Similarity 45.5%; Pred. No. 28;
Matches 10; Conservative 4; Mismatches 7; Indels 1; Gaps 1;
QY 2 KDSHLGAGPAACHGHLILGNP 23
Db 133 RNDNGYCPGPG-HGLMAGNP 153

RESULT 13
AAG28063
ID AAG28063 standard; Protein; 270 AA.
XX AC AAG28063;
XX DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 33140.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence.
XX OS Arabidopsis thaliana.
XX

PN EF1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX PR 25-FEB-1999; 99US-0121825.
PR 03-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 24-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.


```
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 18-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147935.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 27-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
```

```
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 06-OCT-1999; 99US-0157753.
PR 07-OCT-1999; 99US-0157865.
PR 08-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
```

Query Match 32.0%; Score 52.5; DB 21; Length 270;
Best Local Similarity 45.5%; Pred. No. 28;
Matches 10; Conservative 4; Mismatches 7; Indels 1; Gaps 1;

QY 2 KDSHLGAGFAACHGHLILGNP 23
::: ||| |||||
Db 138 RNDNNGYPPGG-HGHLMAGNP 158

RESULT 14

AAG28062
ID AAG28062 standard; Protein; 478 AA.

XX AC AAG28062;

XX DT 17-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 33139.

XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence.

XX OS Arabidopsis thaliana.

XX PN EPI033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134270.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.

| | | |
|----|--------------|---------------|
| PR | 05-OCT-1999; | 9905-0157753. |
| PR | 06-OCT-1999; | 9905-0157865. |
| PR | 07-OCT-1999; | 9905-0158029. |
| PR | 08-OCT-1999; | 9905-0158369. |
| PR | 12-OCT-1999; | 9905-0158362. |
| PR | 13-OCT-1999; | 9905-0159293. |
| PR | 13-OCT-1999; | 9905-0159294. |
| PR | 13-OCT-1999; | 9905-0159295. |
| PR | 14-OCT-1999; | 9905-0159329. |
| PR | 14-OCT-1999; | 9905-0159330. |
| PR | 14-OCT-1999; | 9905-0159331. |
| PR | 14-OCT-1999; | 9905-0159637. |
| PR | 14-OCT-1999; | 9905-0159638. |
| PR | 18-OCT-1999; | 9905-0159584. |
| PR | 21-OCT-1999; | 9905-0160741. |
| PR | 21-OCT-1999; | 9905-0160767. |
| PR | 21-OCT-1999; | 9905-0160768. |
| PR | 21-OCT-1999; | 9905-0160770. |
| PR | 21-OCT-1999; | 9905-0160814. |
| PR | 21-OCT-1999; | 9905-0160815. |
| PR | 22-OCT-1999; | 9905-0160980. |
| PR | 22-OCT-1999; | 9905-0160981. |
| PR | 22-OCT-1999; | 9905-0160989. |
| PR | 25-OCT-1999; | 9905-0161404. |
| PR | 25-OCT-1999; | 9905-0161405. |
| PR | 25-OCT-1999; | 9905-0161406. |
| PR | 26-OCT-1999; | 9905-0161359. |
| PR | 26-OCT-1999; | 9905-0161360. |
| PR | 26-OCT-1999; | 9905-0161361. |
| PR | 28-OCT-1999; | 9905-0161920. |
| PR | 28-OCT-1999; | 9905-0161992. |
| PR | 28-OCT-1999; | 9905-0161993. |
| PR | 29-OCT-1999; | 9905-0162142. |

| | |
|-----------|--|
| RESULT 15 | |
| ABG21033 | |
| ID | ABG21033 standard; Protein; 193 AA. |
| XX | |
| AC | ABG21033; |
| XX | |
| DT | 18-FEB-2002 (first entry) |
| XX | |
| DE | Novel human diagnostic protein #21024. |
| XX | |
| KN | Human; chromosome mapping; gene mapping; gene therapy; forensic; |
| KW | food supplement; medical imaging; diagnostic; genetic disorder. |
| XX | |
| OS | Homo sapiens. |
| XX | |
| PN | WO200175067-A2. |
| XX | |
| PD | 11-OCT-2001. |
| XX | |
| PF | 30-MAR-2001; 2001WO-US08631. |
| XX | |
| PR | 31-MAR-2000; 2000US-0540217. |
| PR | 23-AUG-2000; 2000US-0649167. |
| XX | |
| PA | (HYSE-) HYSBQ INC. |
| XX | |
| PI | Dmanac RT, Liu C, Tang YT; |
| XX | |
| DR | WPI; 2001-639362/73. |
| DR | N-PSDB; AAS85220. |

| | |
|----|---|
| XX | New isolated polynucleotide and encoded polypeptides, useful in |
| PT | diagnostics, forensics, gene mapping, identification of mutations |
| PT | responsible for genetic disorders or other traits and to assess |
| PT | biodiversity - |
| XX | |
| XX | Claim 20; SEQ ID No 51392; 103pp; English. |
| XX | |
| CC | The invention relates to isolated polynucleotide (I) and |
| CC | polypeptide (II) sequences. (I) is useful as hybridisation probes, |
| CC | polymerase chain reaction (PCR) primers, oligomers, and for chromosome |
| CC | and gene mapping, and in recombinant production of (II). The |
| CC | polynucleotides are also used in diagnostics as expressed sequence tags |
| CC | for identifying expressed genes. (I) is useful in gene therapy techniques |
| CC | to restore normal activity of (II) or to treat disease states involving |
| CC | (II). (II) is useful for generating antibodies against it, detecting and |
| CC | quantitating a polypeptide in tissue, as molecular weight markers and as |
| CC | a food supplement. (II) and its binding partners are useful in medical |
| CC | imaging of sites expressing (II). (I) and (II) are useful for treating |
| CC | disorders involving aberrant protein expression or biological activity. |
| CC | The polypeptide and polynucleotide sequences have applications in |
| CC | diagnostics, forensics, gene mapping, identification of mutations |
| CC | responsible for genetic disorders or other traits to assess biodiversity |
| CC | and to produce other types of data and products dependent on DNA and |
| CC | amino acid sequences. ABG00010-ABG30377 represent novel human |
| CC | diagnostic amino acid sequences of the invention. |
| CC | Note: The sequence data for this patent did not appear in the printed |
| CC | specification, but was obtained in electronic format directly from WIPO |
| CC | at ftp.wipo.int/pub/published_pct_sequences. |
| XX | |
| XX | Sequence 193 AA; |
| XX | |
| CC | Query Match 31.1%; Score 51; DB 22; Length 193; |
| CC | Best Local Similarity 45.8%; Pred. No. 31; |
| CC | Matches 11; Conservative 1; Mismatches 6; Indels 6; Gaps 1 |
| QY | 3 DSHLGAGPAAACHGHLILGNPKNS 26 |
| | |
| | |
| Db | 55 DSHLGSGPACTA-----GGPRTS 72 |

Search completed: June 3, 2003, 19:22:35
Job time : 8.93902 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2003, 19:24:01 ; Search time 51.565 Seconds
(without alignments)
422.051 Million cell updates/sec

Title: US-09-938-703-6

Perfect score: 1122
Sequence: 1 MDQVSSPIYDINTYSEPC.....AACHGHLGNPKNSASYSK 215

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 303519 seqs, 101223694 residues

Total number of hits satisfying chosen parameters: 383519

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/FCI_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Query Score | Match | Length | ID | Description |
|------------|-------------|-------|--------|----|-------------------|
| 1 | 1122 | 100.0 | 215 | 10 | US-09-938-719-6 |
| 2 | 1122 | 100.0 | 215 | 10 | US-09-939-226-6 |
| 3 | 1122 | 100.0 | 215 | 10 | US-09-938-703-6 |
| 4 | 958 | 85.4 | 184 | 10 | US-09-938-719-4 |
| 5 | 958 | 85.4 | 184 | 10 | US-09-939-226-4 |
| 6 | 958 | 85.4 | 184 | 10 | US-09-938-703-4 |
| 7 | 958 | 85.4 | 352 | 9 | US-10-232-686-2 |
| 8 | 958 | 85.4 | 352 | 9 | US-10-086-814-1 |
| 9 | 958 | 85.4 | 352 | 9 | US-09-734-221A-14 |
| 10 | 958 | 85.4 | 352 | 10 | US-09-725-285-2 |
| 11 | 958 | 85.4 | 352 | 10 | US-09-759-841-2 |
| 12 | 958 | 85.4 | 352 | 10 | US-09-779-879A-22 |
| 13 | 958 | 85.4 | 352 | 10 | US-09-779-880A-22 |
| 14 | 958 | 85.4 | 352 | 10 | US-09-813-653-15 |
| 15 | 958 | 85.4 | 352 | 10 | US-09-796-202-1 |
| 16 | 958 | 85.4 | 352 | 10 | US-09-195-662A-2 |
| 17 | 958 | 85.4 | 352 | 10 | US-09-339-912A-2 |
| 18 | 958 | 85.4 | 352 | 10 | US-09-938-719-5 |
| 19 | 958 | 85.4 | 352 | 10 | US-09-939-226-5 |

| | | | | | | |
|----|-------|------|-----|----|-------------------|--------------------|
| 20 | 958 | 85.4 | 352 | 10 | US-09-938-703-5 | Sequence 5, Appli |
| 21 | 958 | 85.4 | 352 | 10 | US-09-502-783A-2 | Sequence 2, Appli |
| 22 | 958 | 85.4 | 352 | 12 | US-10-106-623-2 | Sequence 2, Appli |
| 23 | 952 | 84.8 | 352 | 10 | US-09-813-653-17 | Sequence 17, Appli |
| 24 | 943 | 84.0 | 352 | 10 | US-09-779-879A-2 | Sequence 2, Appli |
| 25 | 943 | 84.0 | 352 | 10 | US-09-779-880A-2 | Sequence 2, Appli |
| 26 | 933 | 83.2 | 352 | 12 | US-10-106-623-20 | Sequence 20, Appli |
| 27 | 695 | 61.9 | 360 | 10 | US-09-131-827A-20 | Sequence 20, Appli |
| 28 | 694 | 61.9 | 344 | 9 | US-10-232-686-9 | Sequence 9, Appli |
| 29 | 694 | 61.9 | 344 | 10 | US-09-779-879A-9 | Sequence 9, Appli |
| 30 | 694 | 61.9 | 344 | 10 | US-09-779-880A-9 | Sequence 9, Appli |
| 31 | 694 | 61.9 | 347 | 10 | US-09-104-792-3 | Sequence 3, Appli |
| 32 | 694 | 61.9 | 360 | 10 | US-09-131-827A-2 | Sequence 2, Appli |
| 33 | 692 | 61.7 | 360 | 10 | US-09-938-719-7 | Sequence 7, Appli |
| 34 | 692 | 61.7 | 360 | 10 | US-09-939-226-7 | Sequence 7, Appli |
| 35 | 692 | 61.7 | 360 | 10 | US-09-938-703-7 | Sequence 7, Appli |
| 36 | 608.5 | 54.2 | 329 | 10 | US-09-725-285-9 | Sequence 9, Appli |
| 37 | 608.5 | 54.2 | 329 | 10 | US-09-195-662A-9 | Sequence 9, Appli |
| 38 | 608.5 | 54.2 | 329 | 10 | US-09-339-912A-9 | Sequence 9, Appli |
| 39 | 608.5 | 54.2 | 329 | 10 | US-09-502-783A-9 | Sequence 9, Appli |
| 40 | 598 | 53.3 | 355 | 10 | US-09-961-068-1 | Sequence 1, Appli |
| 41 | 598 | 53.3 | 355 | 10 | US-09-960-547-1 | Sequence 1, Appli |
| 42 | 598 | 53.3 | 375 | 9 | US-10-219-834-78 | Sequence 78, Appli |
| 43 | 594 | 52.9 | 355 | 10 | US-09-938-719-9 | Sequence 9, Appli |
| 44 | 594 | 52.9 | 355 | 10 | US-09-939-226-9 | Sequence 9, Appli |
| 45 | 594 | 52.9 | 355 | 10 | US-09-938-703-9 | Sequence 9, Appli |

ALIGNMENTS

RESULT 1

US-09-938-719-6

; Sequence 6, Application US/09938719

; Patent No. US20020106742A1

GENERAL INFORMATION:

APPLICANT: SAMSON, MICHEL

PARMENTIER, MARC

VASSART, GILBERT

LIBERT, FREDERICK

TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR

AND NUCLEIC ACID MOLECULES ENCODING SAID RECEPTOR

NUMBER OF SEQUENCES: 17

CORRESPONDENCE ADDRESS:

ADDRESSEE: Knobbe, Martens, Olson & Bear

STREET: 620 Newport Center Drive 16th Floor

CITY: Newport Beach

STATE: CA

COUNTRY: U.S.A.

ZIP: 92660

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

CURRENT APPLICATION NUMBER: US/09/938,719

FILING DATE: 24-Aug-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/626,939

FILING DATE: 27-JULY-2000

ATTORNEY/AGENT INFORMATION:

NAME: Altman, Daniel E

REGISTRATION NUMBER: 34,115

REFERENCE/DOCKET NUMBER: <Unknown>

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 215 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 6:

US-09-938-719-6

Query Match 100.0%; Score 1122; DB 10; Length 215;
Best Local Similarity 100.0%; Pred. No. 4.2e-104;
Matches 215; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDYQVSSPIYDINTYSEPCQKINVKQIAARLLPPLYSLVFIQFVGNMVLVILINCKR 60
DB 1 MDYQVSSPIYDINTYSEPCQKINVKQIAARLLPPLYSLVFIQFVGNMVLVILINCKR 60
QY 61 LKSMTDIYLLNLAIISDLFLTLVFPFAHYAAQWDFGNTMCOLLTGLYIFGFSGIFII 120
DB 61 LKSMTDIYLLNLAIISDLFLTLVFPFAHYAAQWDFGNTMCOLLTGLYIFGFSGIFII 120
QY 121 LITDRYLAVVHAFKARTVTFGVVTSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
DB 121 LITDRYLAVVHAFKARTVTFGVVTSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
QY 181 HPPIKSHLGAGPAAACHGHLILGNPKNSASVSK 215
DB 181 HPPIKSHLGAGPAAACHGHLILGNPKNSASVSK 215

RESULT 2

US-09-939-226-6
; Sequence 6, Application US/09939226
; Patent No. US20020110805A1
; GENERAL INFORMATION:
; APPLICANT: SAMSON, MICHEL
; PARMENTIER, MARC
; VASSART, GILBERT
; LIBERT, FREDERICK
; TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR
; AND NUCLEIC ACID MOLECULES ENCODING SAID RECEPTOR
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/939,226
; FILING DATE: 24-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/626,939
; FILING DATE: 2000-07-27
; ATTORNEY/AGENT INFORMATION:
; NAME: Altman, Daniel E
; REGISTRATION NUMBER: 34,115
; REFERENCE/DOCKET NUMBER: <Unknown>
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 215 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-939-226-6

Query Match 100.0%; Score 1122; DB 10; Length 215;
Best Local Similarity 100.0%; Pred. No. 4.2e-104;
Matches 215; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDYQVSSPIYDINTYSEPCQKINVKQIAARLLPPLYSLVFIQFVGNMVLVILINCKR 60

DB 1 MDYQVSSPIYDINTYSEPCQKINVKQIAARLLPPLYSLVFIQFVGNMVLVILINCKR 60
QY 61 LKSMTDIYLLNLAIISDLFLTLVFPFAHYAAQWDFGNTMCOLLTGLYIFGFSGIFII 120
DB 61 LKSMTDIYLLNLAIISDLFLTLVFPFAHYAAQWDFGNTMCOLLTGLYIFGFSGIFII 120
QY 121 LITDRYLAVVHAFKARTVTFGVVTSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
DB 121 LITDRYLAVVHAFKARTVTFGVVTSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
QY 181 HPPIKSHLGAGPAAACHGHLILGNPKNSASVSK 215
DB 181 HPPIKSHLGAGPAAACHGHLILGNPKNSASVSK 215

RESULT 3

US-09-938-703-6
; Sequence 6, Application US/09938703
; Patent No. US20020110870A1
; GENERAL INFORMATION:
; APPLICANT: SAMSON, MICHEL
; PARMENTIER, MARC
; VASSART, GILBERT
; LIBERT, FREDERICK
; TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR
; AND NUCLEIC ACID MOLECULES ENCODING SAID RECEPTOR
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/938,703
; FILING DATE: 24-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/626,939
; FILING DATE: 2000-07-27
; ATTORNEY/AGENT INFORMATION:
; NAME: Altman, Daniel E
; REGISTRATION NUMBER: 34,115
; REFERENCE/DOCKET NUMBER: <Unknown>
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 215 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-938-703-6

Query Match 100.0%; Score 1122; DB 10; Length 215;
Best Local Similarity 100.0%; Pred. No. 4.2e-104;
Matches 215; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDYQVSSPIYDINTYSEPCQKINVKQIAARLLPPLYSLVFIQFVGNMVLVILINCKR 60
DB 1 MDYQVSSPIYDINTYSEPCQKINVKQIAARLLPPLYSLVFIQFVGNMVLVILINCKR 60
QY 61 LKSMTDIYLLNLAIISDLFLTLVFPFAHYAAQWDFGNTMCOLLTGLYIFGFSGIFII 120
DB 61 LKSMTDIYLLNLAIISDLFLTLVFPFAHYAAQWDFGNTMCOLLTGLYIFGFSGIFII 120
QY 121 LITDRYLAVVHAFKARTVTFGVVTSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180

Db 121 LTTIDRYLVAVHAFKARTVTFGVVTSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
QY 181 HFPYKDSHLGAGPAAACHGHLILGNPKNSASVSK 215
Db 181 HFPYKDSHLGAGPAAACHGHLILGNPKNSASVSK 215

RESULT 4

US-09-938-719-4
; Sequence 4, Application US/09938719
; Patent No. US20020106742A1
; GENERAL INFORMATION:
; APPLICANT: SAMSON, MICHEL
; PARMENTIER, MARC
; VASSART, GILBERT
; LIBERT, FREDERICK
; TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR
; AND NUCLEIC ACID MOLECULES ENCODING SAID RECEPTOR
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/938,719
; FILING DATE: 24-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/626,939
; FILING DATE: 27-JULY-2000
; ATTORNEY/AGENT INFORMATION:
; NAME: Altman, Daniel E
; REGISTRATION NUMBER: 34,115
; REFERENCE/DOCKET NUMBER: <Unknown>
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 184 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-938-719-4

Query Match 85.4%; Score 958; DB 10; Length 184;
Best Local Similarity 100.0%; Pred. No. 7.5e-86;
Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDYQVSSPIYDINTYTSEPCQKINVKQIAARLLPPLYSLVFIFGFGVGNMLVILINCKR 60
Db 1 MDYQVSSPIYDINTYTSEPCQKINVKQIAARLLPPLYSLVFIFGFGVGNMLVILINCKR 60
QY 61 LKSMTDIYLLNLAISDLFFLLTPFWAHYAAQWDFGNTMCQLLTGLYIFGFFSGIFFII 120
Db 61 LKSMTDIYLLNLAISDLFFLLTPFWAHYAAQWDFGNTMCQLLTGLYIFGFFSGIFFII 120
QY 121 LTTIDRYLVAVHAFKARTVTFGVVTSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
Db 121 LTTIDRYLVAVHAFKARTVTFGVVTSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
QY 181 HFPY 184
Db 181 HFPY 184

RESULT 5

US-09-939-226-4
; Sequence 4, Application US/09939226
; Patent No. US20020110805A1
; GENERAL INFORMATION:
; APPLICANT: SAMSON, MICHEL
; PARMENTIER, MARC
; VASSART, GILBERT
; LIBERT, FREDERICK
; TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR
; AND NUCLEIC ACID MOLECULES ENCODING SAID RECEPTOR
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/939,226
; FILING DATE: 24-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/626,939
; FILING DATE: 2000-07-27
; ATTORNEY/AGENT INFORMATION:
; NAME: Altman, Daniel E
; REGISTRATION NUMBER: 34,115
; REFERENCE/DOCKET NUMBER: <Unknown>
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 184 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-939-226-4

Query Match 85.4%; Score 958; DB 10; Length 184;
Best Local Similarity 100.0%; Pred. No. 7.5e-88;
Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDYQVSSPIYDINTYTSEPCQKINVKQIAARLLPPLYSLVFIFGFGVGNMLVILINCKR 60
Db 1 MDYQVSSPIYDINTYTSEPCQKINVKQIAARLLPPLYSLVFIFGFGVGNMLVILINCKR 60
QY 61 LKSMTDIYLLNLAISDLFFLLTPFWAHYAAQWDFGNTMCQLLTGLYIFGFFSGIFFII 120
Db 61 LKSMTDIYLLNLAISDLFFLLTPFWAHYAAQWDFGNTMCQLLTGLYIFGFFSGIFFII 120
QY 121 LTTIDRYLVAVHAFKARTVTFGVVTSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
Db 121 LTTIDRYLVAVHAFKARTVTFGVVTSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
QY 181 HFPY 184
Db 181 HFPY 184

RESULT 6

US-09-938-703-4
; Sequence 4, Application US/09938703
; Patent No. US20020110870A1
; GENERAL INFORMATION:
; APPLICANT: SAMSON, MICHEL
; PARMENTIER, MARC
; VASSART, GILBERT
; LIBERT, FREDERICK

;/ TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR
;/ AND NUCLEIC ACID MOLECULES ENCODING SAID RECEPTOR
;/ NUMBER OF SEQUENCES: 17
;/ CORRESPONDENCE ADDRESS:
;/ ADDRESSEE: Knobbe, Martens, Olson & Bear
;/ STREET: 620 Newport Center Drive 16th Floor
;/ City: Newport Beach
;/ STATE: CA
;/ COUNTRY: U.S.A.
;/ ZIP: 92660
;/ COMPUTER READABLE FORM:
;/ MEDIUM TYPE: Floppy disk
;/ COMPUTER: IBM PC compatible
;/ OPERATING SYSTEM: PC-DOS/MS-DOS
;/ SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
;/ CURRENT APPLICATION DATA:
;/ APPLICATION NUMBER: US/09/938,703
;/ FILING DATE: 24-Aug-2001
;/ CLASSIFICATION: <Unknown>
;/ PRIOR APPLICATION DATA:
;/ APPLICATION NUMBER: 09/626,939
;/ FILING DATE: 2000-07-27
;/ ATTORNEY/AGENT INFORMATION:
;/ NAME: Altman, Daniel E
;/ REGISTRATION NUMBER: 34,115
;/ REFERENCE/DOCKET NUMBER: <Unknown>
;/ INFORMATION FOR SEQ ID NO: 4:
;/ SEQUENCE CHARACTERISTICS:
;/ LENGTH: 184 amino acids
;/ TYPE: amino acid
;/ TOPOLOGY: linear
;/ MOLECULE TYPE: protein
;/ SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-938-703-4

Query Match 85.4%; Score 958; DB 10; Length 184;
Best Local Similarity 100.0%; Pred. No. 7.5e-88;
Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDYQVSSPIYDINVTSEPCQKINVKQIAARLLPPLYSLVIFGFGNMLVILLINCKR 60
Db 1 MDYQVSSPIYDINVTSEPCQKINVKQIAARLLPPLYSLVIFGFGNMLVILLINCKR 60
QY 61 LKSMTDIYLLNLAIISDLFLLTPVFWAHYAAQWDFGNTMCLLTGLYFGFSGIFII 120
Db 61 LKSMTDIYLLNLAIISDLFLLTPVFWAHYAAQWDFGNTMCLLTGLYFGFSGIFII 120
QY 121 LLTIDRYLAVVHAFVAFKARTVTFGVTSVITWVAVFASLPGLIIFTRSQKGLHYTCSS 180
Db 121 LLTIDRYLAVVHAFVAFKARTVTFGVTSVITWVAVFASLPGLIIFTRSQKGLHYTCSS 180
QY 181 HFPY 184
Db 181 HFPY 184

RESULT 7
US-10-232-686-2
;/ Sequence 2, Application US/10232686
;/ Publication No. US20030023044A1
;/ GENERAL INFORMATION:
;/ APPLICANT: Li, Yi
;/ APPLICANT: Ruben, Steven M.
;/ TITLE OF INVENTION: Human G-Protein Chemokine Receptor (CCR5) HDGNR10
;/ FILE REFERENCE: 1488.115000N
;/ CURRENT APPLICATION NUMBER: US/10/232,686
;/ CURRENT FILING DATE: 2002-09-03
;/ PRIOR APPLICATION NUMBER: 09/339,912
;/ PRIOR FILING DATE: 1999-06-25
;/ PRIOR APPLICATION NUMBER: 09/195,662
;/ PRIOR FILING DATE: 1998-11-18
;/ PRIOR APPLICATION NUMBER: 08/466,343
;/ PRIOR FILING DATE: 1995-06-06

;/ NUMBER OF SEQ ID NOS: 9
;/ SOFTWARE: PatentIn version 3.0
;/ SEQ ID NO 2
;/ LENGTH: 352
;/ TYPE: PRT
;/ ORGANISM: Homo sapiens
US-10-232-686-2
Query Match 85.4%; Score 958; DB 9; Length 352;
Best Local Similarity 100.0%; Pred. No. 1.6e-87;
Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDYQVSSPIYDINVTSEPCQKINVKQIAARLLPPLYSLVIFGFGNMLVILLINCKR 60
Db 1 MDYQVSSPIYDINVTSEPCQKINVKQIAARLLPPLYSLVIFGFGNMLVILLINCKR 60
QY 61 LKSMTDIYLLNLAIISDLFLLTPVFWAHYAAQWDFGNTMCLLTGLYFGFSGIFII 120
Db 61 LKSMTDIYLLNLAIISDLFLLTPVFWAHYAAQWDFGNTMCLLTGLYFGFSGIFII 120
QY 121 LLTIDRYLAVVHAFVAFKARTVTFGVTSVITWVAVFASLPGLIIFTRSQKGLHYTCSS 180
Db 121 LLTIDRYLAVVHAFVAFKARTVTFGVTSVITWVAVFASLPGLIIFTRSQKGLHYTCSS 180
QY 181 HFPY 184
Db 181 HFPY 184
RESULT 8
US-10-086-814-1
;/ Sequence 1, Application US/10086814
;/ Publication No. US20030092632A1
;/ GENERAL INFORMATION:
;/ APPLICANT: Dragic, Tatjana
;/ APPLICANT: Olson, William C.
;/ TITLE OF INVENTION: SULFATED CCR5 PEPTIDES FOR HIV-1 INFECTION
;/ FILE REFERENCE: 61010-AB-1
;/ CURRENT APPLICATION NUMBER: US/10/086,814
;/ CURRENT FILING DATE: 2002-02-28
;/ NUMBER OF SEQ ID NOS: 38
;/ SOFTWARE: PatentIn version 3.1
;/ SEQ ID NO 1
;/ LENGTH: 352
;/ TYPE: PRT
;/ ORGANISM: Homo sapiens
US-10-086-814-1
Query Match 85.4%; Score 958; DB 9; Length 352;
Best Local Similarity 100.0%; Pred. No. 1.6e-87;
Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDYQVSSPIYDINVTSEPCQKINVKQIAARLLPPLYSLVIFGFGNMLVILLINCKR 60
Db 1 MDYQVSSPIYDINVTSEPCQKINVKQIAARLLPPLYSLVIFGFGNMLVILLINCKR 60
QY 61 LKSMTDIYLLNLAIISDLFLLTPVFWAHYAAQWDFGNTMCLLTGLYFGFSGIFII 120
Db 61 LKSMTDIYLLNLAIISDLFLLTPVFWAHYAAQWDFGNTMCLLTGLYFGFSGIFII 120
QY 121 LLTIDRYLAVVHAFVAFKARTVTFGVTSVITWVAVFASLPGLIIFTRSQKGLHYTCSS 180
Db 121 LLTIDRYLAVVHAFVAFKARTVTFGVTSVITWVAVFASLPGLIIFTRSQKGLHYTCSS 180
QY 181 HFPY 184
Db 181 HFPY 184

RESULT 9
US-09-734-221A-14
;/ Sequence 14, Application US/09734221A
;/ Publication No. US20030096221A1

GENERAL INFORMATION:
APPLICANT: LITTMAN, DAN R.
DENG, HONGKUI
BILMEIER, WILFRIED
LANDAU, NATHANIEL R.
LIU, RONG
TITLE OF INVENTION: G-COUPLED RECEPTORS ASSOCIATED WITH
MACROPHAGE-TROPIC HIV, AND DIAGNOSTIC AND THERAPEUTIC
USES THEREOF
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: David A. Jackson, Esq.
STREET: 411 Hackensack Ave, Continental Plaza, 4th
Floor
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/734,221A
FILING DATE: 11-Dec-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/666,020
FILING DATE: 19-JUN-1996
APPLICATION NUMBER: US 08/227,319
FILING DATE: 13-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 1049-1-004 N2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 352 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-09-734-221A-14

Query Match 85.4%; Score 958; DB 9; Length 352;
Best Local Similarity 100.0%; Pred. No. 1.6e-87;
Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDYQVSSPIYDINYYTSEPCKINVKQIAARLLPPLYSLVFIFGVGNMLVILLINCKR 60
DB 1 MDYQVSSPIYDINYYTSEPCKINVKQIAARLLPPLYSLVFIFGVGNMLVILLINCKR 60
QY 61 LKSMTDIYLLNLAIISDLFFLLTPVFWAHYAAQWDFGNTMCQLLGLYFIFGFSGIFII 120
DB 61 LKSMTDIYLLNLAIISDLFFLLTPVFWAHYAAQWDFGNTMCQLLGLYFIFGFSGIFII 120
QY 121 LLTIDRYLAVHVPALKARTVFGVTVSVITWVAVFASLPGIIFRSQKEGLHYTCSS 180
DB 121 LLTIDRYLAVHVPALKARTVFGVTVSVITWVAVFASLPGIIFRSQKEGLHYTCSS 180
QY 181 HPFY 184
DB 181 HPFY 184

RESULT 10
US-09-725-285-2
Sequence 2, Application US/09725285
Patent No. US20010000241A1
GENERAL INFORMATION:
APPLICANT: Li, Yi
APPLICANT: Ruben, Steven, M.
TITLE OF INVENTION: Antibodies to Human G-Protein Chemokine Receptor HDGNR10
TITLE OF INVENTION: (CCR5 Receptor)
FILE REFERENCE: 1488.1150003
CURRENT APPLICATION NUMBER: US/09/725,285
CURRENT FILING DATE: 2000-11-29
PRIOR APPLICATION NUMBER: 09/339,912
PRIOR FILING DATE: 1999-06-25
PRIOR APPLICATION NUMBER: 09/195,662
PRIOR FILING DATE: 1998-11-18
PRIOR APPLICATION NUMBER: 08/466,343
PRIOR FILING DATE: 1995-06-06
NUMBER OF SEQ ID NOS: 9
SOFTWARE: PatentIn version 3.0
SEQ ID NO 2
LENGTH: 352
TYPE: PRT
ORGANISM: Artificial Sequence: Genomic
FEATURE:
OTHER INFORMATION: Deduced Amino Acid Sequence
US-09-725-285-2

Query Match 85.4%; Score 958; DB 10; Length 352;
Best Local Similarity 100.0%; Pred. No. 1.6e-87;
Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDYQVSSPIYDINYYTSEPCKINVKQIAARLLPPLYSLVFIFGVGNMLVILLINCKR 60
DB 1 MDYQVSSPIYDINYYTSEPCKINVKQIAARLLPPLYSLVFIFGVGNMLVILLINCKR 60
QY 61 LKSMTDIYLLNLAIISDLFFLLTPVFWAHYAAQWDFGNTMCQLLGLYFIFGFSGIFII 120
DB 61 LKSMTDIYLLNLAIISDLFFLLTPVFWAHYAAQWDFGNTMCQLLGLYFIFGFSGIFII 120
QY 121 LLTIDRYLAVHVPALKARTVFGVTVSVITWVAVFASLPGIIFRSQKEGLHYTCSS 180
DB 121 LLTIDRYLAVHVPALKARTVFGVTVSVITWVAVFASLPGIIFRSQKEGLHYTCSS 180
QY 181 HPFY 184
DB 181 HPFY 184

RESULT 11
US-09-759-841-2
Sequence 2, Application US/09759841
Patent No. US20010039026A1
GENERAL INFORMATION:
APPLICANT: Rickett, Graham A
APPLICANT: Dobbs, Susan
APPLICANT: Perros, Manoussos
TITLE OF INVENTION: Assay Method
FILE REFERENCE: PC10348APME
CURRENT APPLICATION NUMBER: US/09/759,841
CURRENT FILING DATE: 2001-01-12
PRIOR APPLICATION NUMBER: GB 0000661.9
PRIOR FILING DATE: 2000-01-12
PRIOR APPLICATION NUMBER: GB 0000663.5
PRIOR FILING DATE: 2000-01-12
PRIOR APPLICATION NUMBER: GB 0000659.3
PRIOR FILING DATE: 2000-01-12
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 352
TYPE: PRT
ORGANISM: Homo sapiens

US-09-779-880A-22

```

Query Match      85.4%; Score 958; DB 10; Length 352;
Best Local Similarity 100.0%; Pred. No. 1.6e-87;
Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MDYVSSPYDINNYTSEPCOKINVKQIARLLPLYSLVFIQFGVGNMLVILLINCKR 60
      |||
Db       1 MDYVSSPYDINNYTSEPCOKINVKQIARLLPLYSLVFIQFGVGNMLVILLINCKR 60

QY      61 LKSMTDIYLLMLAISDLFFLTVPFWAHAAQDFGNMTCQLLTGLTFLYFGFFSGIFPII 120
      |||
Db       61 LKSMTDIYLLMLAISDLFFLTVPFWAHAAQDFGNMTCQLLTGLTFLYFGFFSGIFPII 120

QY      121 LTTIDRYLVAVHVAEFAKARTVTGGVTVSVITWVAVFASLPGIIFTSQKEGLHYTCSS 180
      |||
Db       121 LTTIDRYLVAVHVAEFAKARTVTGGVTVSVITWVAVFASLPGIIFTSQKEGLHYTCSS 180

QY      181 HPFY 184
      ||||
Db       181 HPFY 184

```

Search completed: June 3, 2003, 19:35:59
Job time : 51.565 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2003, 19:21:31 ; Search time 4.53659 Seconds
(without alignments)
201.057 Million cell updates/sec

Title: US-09-938-703-6_COPY_185_215
Sequence: 1 IKDHLGAGPAAACHGILLGNPKNSASVSK 31

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 2942292 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents.AA:*
1: /cgn2_6/ptodata/1/iaa/5A.COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B.COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A.COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B.COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS.COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------------|
| 1 | 164 | 100.0 | 34 | 4 | US-08-833-752-13 |
| 2 | 164 | 100.0 | 215 | 4 | US-09-087-232A-17 |
| 3 | 164 | 100.0 | 215 | 4 | US-08-833-752-6 |
| 4 | 50 | 30.5 | 501 | 2 | US-08-660-963-13 |
| 5 | 47.5 | 29.0 | 2304 | 4 | US-09-324-867-4 |
| 6 | 47.5 | 29.0 | 2319 | 1 | US-08-212-133A-8 |
| 7 | 47.5 | 29.0 | 2319 | 1 | US-08-474-503-6 |
| 8 | 47.5 | 29.0 | 2319 | 2 | US-08-670-707A-6 |
| 9 | 47.5 | 29.0 | 2319 | 4 | US-09-037-601-6 |
| 10 | 47.5 | 29.0 | 2319 | 4 | US-09-315-179-6 |
| 11 | 47.5 | 29.0 | 2319 | 4 | US-09-523-656-28 |
| 12 | 47.5 | 29.0 | 2319 | 5 | PCT-US94-13200-6 |
| 13 | 45.5 | 27.7 | 1436 | 4 | US-09-209-316-1 |
| 14 | 45.5 | 27.7 | 1471 | 1 | US-08-683-839B-3 |
| 15 | 45.5 | 27.7 | 1661 | 2 | US-08-882-083-2 |
| 16 | 45.5 | 27.7 | 1661 | 2 | US-08-558-107-2 |
| 17 | 45.5 | 27.7 | 1661 | 4 | US-09-243-539-2 |
| 18 | 45.5 | 27.7 | 1958 | 1 | US-07-945-283-2 |
| 19 | 45.5 | 27.7 | 2332 | 1 | US-07-864-004B-4 |
| 20 | 45.5 | 27.7 | 2332 | 1 | US-08-251-937A-4 |
| 21 | 45.5 | 27.7 | 2332 | 1 | US-08-212-133A-2 |
| 22 | 45.5 | 27.7 | 2332 | 1 | US-08-276-594A-2 |
| 23 | 45.5 | 27.7 | 2332 | 1 | US-08-474-503-2 |
| 24 | 45.5 | 27.7 | 2332 | 2 | US-08-670-707A-2 |
| 25 | 45.5 | 27.7 | 2332 | 4 | US-09-037-601-2 |
| 26 | 45.5 | 27.7 | 2332 | 4 | US-09-324-867-3 |
| 27 | 45.5 | 27.7 | 2332 | 4 | US-09-315-179-2 |

28 45.5 27.7 2332 4 US-09-523-656-2 Sequence 2, Appli
29 45.5 27.7 2332 5 PCT-US93-03275-4 Sequence 4, Appli
30 45.5 27.7 2332 5 PCT-US94-13200-2 Sequence 2, Appli
31 45.5 27.7 2351 1 US-08-121-202-2 Sequence 2, Appli
32 45.5 27.7 2351 1 US-08-366-851A-2 Sequence 2, Appli
33 45.5 27.7 2351 6 5171844-2 Patent No. 5171844
34 45.5 27.7 2351 6 5422260-1 Patent No. 5422260
35 45 27.4 272 4 US-09-177-165A-27 Sequence 24, Appli
36 45 27.4 815 4 US-09-177-165A-24 Sequence 24, Appli
37 44.5 27.1 166 3 US-08-513-974B-312 Sequence 312, App
38 44.5 27.1 2232 4 US-09-091-219-25 Sequence 25, Appli
39 44.5 27.1 2247 4 US-09-091-219-2 Sequence 2, Appli
40 44 26.8 188 2 US-08-933-750C-8 Sequence 8, Appli
41 44 26.8 188 4 US-09-234-613-8 Sequence 8, Appli
42 44 26.8 455 2 US-08-870-827-3 Sequence 3, Appli
43 44 26.8 455 4 US-09-317-179-3 Sequence 3, Appli
44 43.5 26.5 782 4 US-09-543-084A-32 Sequence 32, Appli
45 43.5 26.5 782 4 US-09-543-084A-33 Sequence 33, Appli

ALIGNMENTS

RESULT 1
US-08-833-752-13
; Sequence 13, Application US/08833752
; Patent No. 6448375
; GENERAL INFORMATION:
; APPLICANT: SAMSON, MICHEL
; APPLICANT: PARMENTIER, MARC
; APPLICANT: VASSART, GILBERT
; APPLICANT: LIBERT, FREDERICK
; TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR
; TITLE OF INVENTION: AND NUCLEIC ACID MOLECULES ENCODING SAID RECEPTOR
; NUMBER OF INVENTIONS: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/833,752
; FILING DATE: 9-APR-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Altman, Daniel E
; REGISTRATION NUMBER: 34,115
; REFERENCE/DOCKET NUMBER:
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 34 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-833-752-13

Query Match 100.0%; Score 164; DB 4; Length 34;
Best Local Similarity 100.0%; Pred. No. 8.9e-18;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 IKDHLGAGPAAACHGILLGNPKNSASVSK 31
DB 4 IKDHLGAGPAAACHGILLGNPKNSASVSK 34

RESULT 2

```
US-09-087-232A-17
; Sequence 17, Application US/09087232A
; Patent No. 6153431
; GENERAL INFORMATION:
; APPLICANT: Quillent et al.
; TITLE OF INVENTION: HUMAN IMMUNODEFICIENCY VIRUS CO-RECEPTOR
; TITLE OF INVENTION: VARIANTS ASSOCIATED WITH RESISTANCE TO VIRUS INFECTION.
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Baker & Botts, L.L.P. attn. Lisa Kole
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/087,232A
; FILING DATE: 28 MAY 1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/048,057
; FILING DATE: 30 MAY 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: KOLE, LISA B.
; REGISTRATION NUMBER: 35,225
; REFERENCE/DOCKET NUMBER: AP 31115
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 408-2628
; TELEFAX: (212) 765-2519
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 215 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-087-232A-17

Query Match 100.0%; Score 164; DB 4; Length 215;
Best Local Similarity 100.0%; Pred. No. 7.6e-17;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKDHLGAGPAAACHGHLILGNPKNSASYSK 31
Db 185 IKDHLGAGPAAACHGHLILGNPKNSASYSK 215

RESULT 3
US-08-833-752-6
; Sequence 6, Application US/08833752
; Patent No. 6448375
; GENERAL INFORMATION:
; APPLICANT: SAMSON, MICHEL
; APPLICANT: PARMENTIER, MARC
; APPLICANT: VASSART, GILBERT
; APPLICANT: LIBERT, FREDERICK
; TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR
; TITLE OF INVENTION: AND NUCLEIC ACID MOLECULES ENCODING SAID RECEPTOR
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/833,752
; FILING DATE: 9-APR-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Altman, Daniel E
; REGISTRATION NUMBER: 34,115
; REFERENCE/DOCKET NUMBER:
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 215 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-833-752-6

Query Match 100.0%; Score 164; DB 4; Length 215;
Best Local Similarity 100.0%; Pred. No. 7.6e-17;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKDHLGAGPAAACHGHLILGNPKNSASYSK 31
Db 185 IKDHLGAGPAAACHGHLILGNPKNSASYSK 215

RESULT 4
US-08-660-963-13
; Sequence 13, Application US/08660963
; Patent No. 5852187
; GENERAL INFORMATION:
; APPLICANT: Thorne, Michael O.
; APPLICANT: Gaylinn, Bruce D.
; APPLICANT: Horikawa, Reiko
; APPLICANT: Lyons Jr., Charles E.
; TITLE OF INVENTION: MOLECULAR CLONING OF THE OVINE PITUITARY
; TITLE OF INVENTION: GROWTH HORMONE RELEASING HORMONE RECEPTOR
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: POPHAM, HAIR, SCHNOBRICH & KAUFMAN, LTD.
; STREET: Metropolitan Square Building, Suite 800, 1450
; STREET: G. Street
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/660,963
; FILING DATE: 12-JUN-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: O'Shaughnessy, Brian P.
; REGISTRATION NUMBER: 32,747
; REFERENCE/DOCKET NUMBER: 18046.036
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-824-8000
; TELEFAX: 202-824-8199
; TELEX: 248516
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 501 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-660-963-13
```

NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: EMU/76677
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-572-6508
TELEFAX: 404-572-6555
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 2319 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Mus musculus
PUBLICATION INFORMATION:
AUTHORS: Elder, F.
AUTHORS: Lakich, D.
AUTHORS: Gitschier, J.
TITLE: Sequence of the Murine Factor VIII cDNA.
Patent No. 5663060
JOURNAL: Genomics
VOLUME: 16
PAGES: 374-379
DATE: 1993
RELEVANT RESIDUES IN SEQ ID NO: 8: FROM 1 TO 2319
US-08-212-133A-8

Query Match 29.0%; Score 47.5; DB 1; Length 2319;
Best Local Similarity 40.0%; Pred. No. 2.3e+02;
Matches 12; Conservative 3; Mismatches 12; Indels 3; Gaps 2;

QY 2 KDHLG-AGPAACHGHLLGNPKNSASVS 30
: : | | | | | | | | | |
DB 1832 RDMHSLGLPLLCCHANTL--NPAHGQVS 1859

RESULT 7
US-08-474-503-6
Sequence 6, Application US/08474503
Patent No. 574446
GENERAL INFORMATION:
APPLICANT: Emory University
TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kilpatrick & Cody
STREET: 1100 Peachtree Street, Suite 2800
CITY: Atlanta
STATE: Georgia
COUNTRY: US
ZIP: 30309
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,503
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Pratt, John S.
REGISTRATION NUMBER: 29,476
REFERENCE/DOCKET NUMBER: EMU106CIP(3)
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-815-6500
TELEFAX: 404-815-6555
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:

LENGTH: 2319 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ANTI-SENSE: NO
N-terminal
FRAGMENT TYPE:
ORGANISM: Mus musculus
PUBLICATION INFORMATION:
AUTHORS: Elder, F.
AUTHORS: Laskich, D.
AUTHORS: Gitschler, J.
TITLE: Sequence of the Murine Factor VIII cDNA.
Patent No. 574446
JOURNAL: Genomics
VOLUME: 16
PAGES: 374-379
DATE: 1993
RELEVANT RESIDUES IN SEQ ID NO: 6: FROM 1 TO 2319
US-08-474-503-6

Query Match 29.0%; Score 47.5; DB 1; Length 2319;
Best Local Similarity 40.0%; Pred. No. 2.3e+02;
Matches 12; Conservative 3; Mismatches 12; Indels 3; Gaps 2;

QY 2 KDSHLG-AGPAAACHGHLNPKNSASVS 30
DB 1832 RDMHSLGLPILCHANTL--NPAHGRQVS 1859

RESULT 8

US-08-670-707A-6
Sequence 6, Application US/08670707A
Patent No. 5859204

GENERAL INFORMATION:
APPLICANT: Lollar, John S.
TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESS:
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
STREET: 5370 Manhattan Circle Suite 201
CITY: Boulder
STATE: Colorado
COUNTRY: USA
ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/670,707A
FILING DATE: 26-JUN-1996
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US94/13200
FILING DATE: 15-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/212,133
FILING DATE: 11-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/864,004
FILING DATE: 07-APR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Greenlee, Lorraine L.
REGISTRATION NUMBER: 27,894
REFERENCE/DOCKET NUMBER: 75-95F
TELECOMMUNICATION INFORMATION:
TELEPHONE: 303/499-8080
TELEFAX: 303/499-8089
INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:
LENGTH: 2319 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ANTI-SENSE: NO
N-terminal
FRAGMENT TYPE:
ORIGINAL SOURCE:
ORGANISM: Mus musculus
PUBLICATION INFORMATION:
AUTHORS: Elder, F.
AUTHORS: Laskich, D.
AUTHORS: Gitschler, J.
TITLE: Sequence of the Murine Factor VIII cDNA
Patent No. 5859204
JOURNAL: Genomics
VOLUME: 16
PAGES: 374-379
DATE: 1993
RELEVANT RESIDUES IN SEQ ID NO: 6: FROM 1 TO 2319
US-08-670-707A-6

Query Match 29.0%; Score 47.5; DB 2; Length 2319;
Best Local Similarity 40.0%; Pred. No. 2.3e+02;
Matches 12; Conservative 3; Mismatches 12; Indels 3; Gaps 2;

QY 2 KDSHLG-AGPAAACHGHLNPKNSASVS 30
DB 1832 RDMHSLGLPILCHANTL--NPAHGRQVS 1859

RESULT 9

US-09-037-601-6
Sequence 6, Application US/09037601
Patent No. 6180371

GENERAL INFORMATION:
APPLICANT: Lollar, John S.
TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESS:
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
STREET: 5370 Manhattan Circle Suite 201
CITY: Boulder
STATE: Colorado
COUNTRY: USA
ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/037,601
FILING DATE: 26-JUN-1996
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US94/13200
FILING DATE: 15-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/212,133
FILING DATE: 11-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/864,004
FILING DATE: 07-APR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Ferber, Donna M.
REGISTRATION NUMBER: 33,878
REFERENCE/DOCKET NUMBER: 75-95F
TELECOMMUNICATION INFORMATION:
TELEPHONE: 303/499-8080
TELEFAX: 303/499-8089

INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 2319 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Mus musculus
PUBLICATION INFORMATION:
AUTHORS: Elder, F.
AUTHORS: Lickich, D.
AUTHORS: Gitschler, J.
TITLE: Sequence of the Murine Factor VIII cDNA
Patent No. 6180371
JOURNAL: Genomics
VOLUME: 16
PAGES: 374-379
DATE: 1993
RELEVANT RESIDUES IN SEQ ID NO: 6: FROM 1 TO 2319

Query Match 29.0%; Score 47.5; DB 4; Length 2319;
Best Local Similarity 40.0%; Pred. No. 2.3e+02;
Matches 12; Conservative 3; Mismatches 12; Indels 3; Gaps 2;
QY 2 KDSHIG-AGPAAACHGHLILGNPKNSASVS 30
DB 1832 RDMHSLGLLPLLCANTL--NPAHGRQVS 1859

RESULT 10
US-09-315-179-6
Sequence 6, Application US/09315179
Patent No. 6376463
GENERAL INFORMATION:
APPLICANT: Lollar, John S
TITLE OF INVENTION: Modified Factor VIII
FILE REFERENCE: 75-95H
CURRENT APPLICATION NUMBER: US/09/315,179
CURRENT FILING DATE: 1999-05-20
EARLIER APPLICATION NUMBER: U.S. 09/037,601
EARLIER FILING DATE: 1998-03-10
EARLIER APPLICATION NUMBER: U.S. 08/670,707
EARLIER FILING DATE: 1996-06-26
EARLIER APPLICATION NUMBER: PCT/US97/11155
EARLIER FILING DATE: 1997-06-26
EARLIER APPLICATION NUMBER: PCT/US94/13200
EARLIER FILING DATE: 1994-11-15
EARLIER APPLICATION NUMBER: U.S. 08/212,133
EARLIER FILING DATE: 1994-03-11
EARLIER APPLICATION NUMBER: U.S. 07/864,004
EARLIER FILING DATE: 1992-04-07
NUMBER OF SEQ ID NOS: 40
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 6
TYPE: PRT
ORGANISM: Mus musculus
US-09-315-179-6

Query Match 29.0%; Score 47.5; DB 4; Length 2319;
Best Local Similarity 40.0%; Pred. No. 2.3e+02;
Matches 12; Conservative 3; Mismatches 12; Indels 3; Gaps 2;
QY 2 KDSHIG-AGPAAACHGHLILGNPKNSASVS 30
DB 1832 RDMHSLGLLPLLCANTL--NPAHGRQVS 1859

RESULT 11
US-09-523-656-28
Sequence 28, Application US/09523656
Patent No. 6458563
GENERAL INFORMATION:
APPLICANT: Lollar, John
TITLE OF INVENTION: MODIFIED FACTOR VIII
FILE REFERENCE: 75-95I
CURRENT APPLICATION NUMBER: US/09/523,656
CURRENT FILING DATE: 2000-03-10
EARLIER APPLICATION NUMBER: 09/037,601
EARLIER FILING DATE: 1998-03-10
EARLIER APPLICATION NUMBER: 08/670,707
EARLIER FILING DATE: 1996-06-26
NUMBER OF SEQ ID NOS: 38
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 28
LENGTH: 2319
TYPE: PRT
ORGANISM: Mus musculus
US-09-523-656-28

Query Match 29.0%; Score 47.5; DB 4; Length 2319;
Best Local Similarity 40.0%; Pred. No. 2.3e+02;
Matches 12; Conservative 3; Mismatches 12; Indels 3; Gaps 2;
QY 2 KDSHIG-AGPAAACHGHLILGNPKNSASVS 30
DB 1832 RDMHSLGLLPLLCANTL--NPAHGRQVS 1859

RESULT 12
PCT-US94-13200-6
Sequence 6, Application PC/TUS9413200
GENERAL INFORMATION:
APPLICANT: Emory University
TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kilpatrick & Cody
STREET: 1100 Peachtree Street, Suite 2800
CITY: Atlanta
STATE: Georgia
COUNTRY: US
ZIP: 30309
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/13200
FILING DATE: 15-NOV-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: EMU106CIP(2)
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-815-6508
TELEFAX: 404-815-6555
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 2319 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Mus musculus

```

; PUBLICATION INFORMATION:
; AUTHORS: Elder, F.
; AUTHORS: Gitschier, J.
; TITLE: Sequence of the Murine Factor VIII cDNA.
; JOURNAL: Genomics
; VOLUME: 16
; PAGES: 374-379
; DATE: 1993
; RELEVANT RESIDUES IN SEQ ID NO: 6: FROM 1 TO 2319
PCT-US94-13200-6

```

```

Query Match      29.0%; Score 47.5; DB 5; Length 2319;
Best Local Similarity 40.0%; Pred. No. 2.3e+02;
Matches 12; Conservative 3; Mismatches 12; Indels 3; Gaps 2;

```

```

QY      2 KDSHLG-AGPAAACHGHLILGNPKNSASVS 30
Db      1832 RDMHSLGLGPLLICHANTL--NPAHGRQVS 1859

```

```

RESULT 13
US-09-209-916-1
; Sequence 1, Application US/09209916
; Patent No. 6358703
; GENERAL INFORMATION:
; APPLICANT: Cho, Myung-Sam
; APPLICANT: Chan, Sham-yuen
; APPLICANT: Kelsey, William
; APPLICANT: Yee, Helena
; TITLE OF INVENTION: Expression System for Factor VIII
; FILE REFERENCE: MSB-7255
; CURRENT APPLICATION NUMBER: US/09/209,916
; CURRENT FILING DATE: 1998-12-10
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1438
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Derived from
; OTHER INFORMATION: human factor VIII sequence
US-09-209-916-1

```

```

Query Match      27.7%; Score 45.5; DB 4; Length 1438;
Best Local Similarity 40.0%; Pred. No. 2.6e+02;
Matches 12; Conservative 3; Mismatches 12; Indels 3; Gaps 2;

```

```

QY      2 KDSHLG-AGPAAACHGHLILGNPKNSASVS 30
Db      951 KDVSGLIGPLLVCHTNTL--NPAHGRQVT 978

```

```

RESULT 14
US-08-683-839B-3
; Sequence 3, Application US/08683839B
; Patent No. 5744326
; GENERAL INFORMATION:
; APPLICANT: Ill, Charles . R. et al.
; TITLE OF INVENTION: Use of Viral Cis-Acting Post-Transcriptional
; TITLE OF INVENTION: Regulatory Sequences To Increase Expression of
; TITLE OF INVENTION: Introns Genes Containing Near-Consensus Splice Sites
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```

```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/683,839B
; FILING DATE: 11-MARCH-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Remillard, Jane E.
; REGISTRATION NUMBER: 38,872
; REFERENCE/DOCKET NUMBER: TTI-138
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1471 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-683-839B-3

```

```

Query Match      27.7%; Score 45.5; DB 1; Length 1471;
Best Local Similarity 40.0%; Pred. No. 2.7e+02;
Matches 12; Conservative 3; Mismatches 12; Indels 3; Gaps 2;

```

```

QY      2 KDSHLG-AGPAAACHGHLILGNPKNSASVS 30
Db      984 KDVSGLIGPLLVCHTNTL--NPAHGRQVT 1011

```

```

RESULT 15
US-08-882-083-2
; Sequence 2, Application US/08882083
; Patent No. 5869292
; GENERAL INFORMATION:
; APPLICANT: VOORBERG, Johannes J.
; TITLE OF INVENTION: HYBRID PROTEINS WITH MODIFIED ACTIVITY
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/882,083
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/558,107
; FILING DATE: 13-NOV-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ISACSON, John P.
; REGISTRATION NUMBER: 33,715
; REFERENCE/DOCKET NUMBER: 30472/212
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1661 amino acids
; TYPE: amino acid

```



```
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-882-083-2

Query Match      27.7%; Score 45.5; DB 2; Length 1661;
Best Local Similarity 40.0%; Pred. No. 3.1e+02;
Matches 12; Conservative 3; Mismatches 12; Indels 3; Gaps 2;

QY      2 KDSHLG-AGPAAACHGHLGLGNPKNSASVS 30
Db      1174 KDVHSLGLPLLYCHTNTL--NPAHGQVT 1201
```

Search completed: June 3, 2003, 19:25:11
Job time : 5.53659 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 3, 2003, 18:44:09 ; Search time 27.6947 Seconds
(without alignments)
3063.973 Million cell updates/sec

Title: US-09-938-703-3
Perfect score: 2628
Sequence: 1 GAATTCGCCCAACAGACCA.....AGTAGATAGATCCGAATTC 1442

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 262574 seqs, 29422922 residues
Total number of hits satisfying chosen parameters: 525148

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Listing first 45 summaries

Command line parameters:
-MODE=frame_n2p_model -DEV=xlp
-Q=/cgn2.1/USPTO.spool/US09938703/runat_03062003_161408_22885/app_query.fasta_1.1870
-DB=Issued Patents AA -SUFFIX=ra1 -MINMATCH=0.1 -LOOPTC=0
-LOOPT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORW=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09938703_ECGN_1.1.42/runat_03062003_161408_22885 -NCPD=6 -ICPD=3
-NO_MAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGIAG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents AA.*
1: /cgn2.6/ptodata/1/iaa/5A.COMB.pcp.*
2: /cgn2.6/ptodata/1/iaa/5B.COMB.pcp.*
3: /cgn2.6/ptodata/1/iaa/6A.COMB.pcp.*
4: /cgn2.6/ptodata/1/iaa/6B.COMB.pcp.*
5: /cgn2.6/ptodata/1/iaa/PTCUS.COMB.pcp.*
6: /cgn2.6/ptodata/1/iaa/backfiles1.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Query Score | Match | Length | ID | Description |
|------------|-------------|-------|--------|----|-------------------|
| 1 | 1745.5 | 66.4 | 352 | 4 | US-09-087-232A-13 |
| 2 | 1745.5 | 66.4 | 352 | 4 | US-08-861-105-14 |
| 3 | 1745.5 | 66.4 | 352 | 4 | US-08-575-967A-2 |
| 4 | 1745.5 | 66.4 | 352 | 4 | US-08-833-752-5 |
| 5 | 1739.5 | 66.2 | 352 | 4 | US-09-045-583-52 |
| 6 | 1739.5 | 66.2 | 352 | 4 | US-09-534-185-52 |
| 7 | 1730.5 | 65.8 | 352 | 3 | US-08-466-343D-2 |
| 8 | 1718.5 | 65.4 | 352 | 4 | US-09-517-605-5 |
| 9 | 1463.5 | 55.7 | 354 | 4 | US-08-724-984A-2 |
| 10 | 1331 | 50.6 | 347 | 1 | US-08-461-244-3 |
| 11 | 1331 | 50.6 | 360 | 1 | US-08-450-393A-4 |
| 12 | 1331 | 50.6 | 360 | 4 | US-08-446-669-4 |

| | | | | | | |
|----|--------|------|-----|---|-------------------|-------------------|
| 13 | 1331 | 50.6 | 360 | 4 | US-09-045-583-50 | Sequence 50, Appl |
| 14 | 1331 | 50.6 | 360 | 4 | US-09-534-185-50 | Sequence 50, Appl |
| 15 | 1331 | 50.6 | 360 | 5 | PCT-US95-00476-4 | Sequence 4, Appl |
| 16 | 1317 | 50.1 | 360 | 4 | US-08-833-752-7 | Sequence 7, Appl |
| 17 | 1311.5 | 49.9 | 360 | 4 | US-09-045-583-51 | Sequence 51, Appl |
| 18 | 1311.5 | 49.9 | 360 | 4 | US-09-534-185-51 | Sequence 51, Appl |
| 19 | 1191 | 45.3 | 374 | 1 | US-08-450-393A-2 | Sequence 2, Appl |
| 20 | 1191 | 45.3 | 374 | 4 | US-08-446-669-2 | Sequence 2, Appl |
| 21 | 1191 | 45.3 | 374 | 5 | PCT-US95-00476-2 | Sequence 9, Appl |
| 22 | 1190.5 | 45.3 | 344 | 3 | US-08-466-343D-9 | Sequence 2, Appl |
| 23 | 1122 | 42.7 | 215 | 4 | US-09-087-232A-17 | Sequence 17, Appl |
| 24 | 1122 | 42.7 | 215 | 4 | US-08-833-752-6 | Sequence 6, Appl |
| 25 | 1003.5 | 38.2 | 355 | 1 | US-08-012-988A-2 | Sequence 2, Appl |
| 26 | 1003.5 | 38.2 | 355 | 1 | US-08-450-393A-5 | Sequence 5, Appl |
| 27 | 1003.5 | 38.2 | 355 | 4 | US-08-446-669-5 | Sequence 5, Appl |
| 28 | 1003.5 | 38.2 | 355 | 4 | US-09-239-938-1 | Sequence 1, Appl |
| 29 | 1003.5 | 38.2 | 355 | 5 | PCT-US95-00476-5 | Sequence 5, Appl |
| 30 | 976.5 | 37.2 | 355 | 4 | US-08-833-752-9 | Sequence 9, Appl |
| 31 | 958 | 36.5 | 184 | 4 | US-08-833-752-4 | Sequence 4, Appl |
| 32 | 949.5 | 36.1 | 355 | 4 | US-09-045-583-53 | Sequence 53, Appl |
| 33 | 949.5 | 36.1 | 355 | 4 | US-09-534-185-53 | Sequence 4, Appl |
| 34 | 912 | 34.7 | 355 | 4 | US-08-575-967A-4 | Sequence 1, Appl |
| 35 | 912 | 34.7 | 355 | 4 | US-08-847-296B-1 | Sequence 54, Appl |
| 36 | 912 | 34.7 | 355 | 4 | US-09-045-583-54 | Sequence 54, Appl |
| 37 | 891.5 | 33.9 | 360 | 4 | US-09-534-185-54 | Sequence 20, Appl |
| 38 | 891.5 | 33.9 | 360 | 4 | US-08-875-573-20 | Sequence 2, Appl |
| 39 | 891.5 | 33.9 | 360 | 4 | US-09-232-878-2 | Sequence 55, Appl |
| 40 | 891.5 | 33.9 | 360 | 4 | US-09-045-583-55 | Sequence 55, Appl |
| 41 | 891.5 | 33.9 | 360 | 4 | US-09-534-185-55 | Sequence 8, Appl |
| 42 | 873 | 33.2 | 355 | 4 | US-08-833-752-8 | Sequence 10, Appl |
| 43 | 857.5 | 32.6 | 360 | 4 | US-08-833-752-10 | Sequence 2, Appl |
| 44 | 716 | 27.2 | 355 | 1 | US-08-461-244-2 | Sequence 56, Appl |
| 45 | 716 | 27.2 | 355 | 4 | US-09-045-583-56 | |

ALIGNMENTS

RESULT 1
US-09-087-232A-13
; Sequence 13, Application US/09087232A
; Patent No. 6153431
; GENERAL INFORMATION:
; APPLICANT: Quilient et al.
; TITLE OF INVENTION: HUMAN IMMUNODEFICIENCY VIRUS CO-RECEPTOR
; TITLE OF INVENTION: VARIANTS ASSOCIATED WITH RESISTANCE TO VIRUS INFECTION.
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Baker & Botts, L.L.P. attn. Lisa Kole
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/087/232A
; FILING DATE: 28 MAY 1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/048,057
; FILING DATE: 30 MAY 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: KOLE, LISA B.
; REGISTRATION NUMBER: 35,225
; REFERENCE/DOCKET NUMBER: AP 31115
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 408-2628
; TELEFAX: (212) 765-2519
; INFORMATION FOR SEQ ID NO: 13:

SEQUENCE CHARACTERISTICS:
LENGTH: 352 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-087-232A-13

Alignment Scores:

Pred. No.: 5,68e-182 Length: 352
Score: 1745.50 Matches: 340
Percent Similarity: 96.60% Conservative: 1
Best Local Similarity: 96.32% Mismatches: 0
Query Match: 66.42% Indels: 12
DB: 4 Gaps: 1

US-09-938-703-3 (1-1442) x US-09-087-232A-13 (1-352)

QY 240 ATGGATTATCAAGTGTCAAGTCCAAATCTATGACATCAATATTATATACATCGAGCCCTGC 299
Db 1 MetAspTyrGlnValSerSerProIleTyrAspIleAsnTyrThrSerGluProCys 20
QY 300 CAAATAATCAATGTAGCAATCGCAGCCGCTCTGCTCGCTCTACTCTACTGCTGTG 359
Db 21 GlnLysIleAsnValLysGlnIleAlaAlaArgLeuLeuProLeuTyrSerLeuVal 40
QY 360 TTCATCTTGTGTTTGTGGCAACATGCTGGTCATCTCCTCATCTCATCTGTAATCTGAAAGG 419
Db 41 PheIlePheGlyPheValGlyAsnMetLeuValIleLeuIleAsnCysLysArg 60
QY 420 CTGAAGACGATGATGATCTACTGCTGCTCAACCTGGCCATCTCTGACCTGTTTTCCTT 479
Db 61 LeuLysSerMetThrAspIleTyrLeuLeuAsnLeuAlaIleSerAspLeuPheLeu 80
QY 480 CTACTGTCCTCTGGGCTCACTATGCTGGCCGCGCCAGTGGAGCTTTGGAATACATG 539
Db 81 LeuThrValProPheTrpAlaHisTyrAlaAlaAlaGlnTrpAspPheGlyAsnThrMet 100
QY 540 TGTCAACTCTGTACAGGCTCTATTTATAGGCTTCTCTGGAATCTCTTCATCATC 599
Db 101 CysGlnLeuLeuThrGlyLeuTyrPheIleGlyPhePheSerGlyIlePhePheIle 120
QY 600 CTCTGCAATCAATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 659
Db 121 LeuLeuThrIleAspArgTyrLeuAlaValAlaHisAlaValPheAlaLeuLysAlaArg 140
QY 660 ACGTCACTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 719
Db 141 ThrValThrPheGlyValValThrSerValIleThrTrpValValAlaValPheAlaSer 160
QY 720 CTCCAGGAATCATCTTTACCATCTTCAAAAGAGCTTTCATTCACCTGCGAGCTCT 779
Db 161 LeuProGlyIleIlePheThrArgSerGlnLysGlnLysGlnLysTyrThrCysSerSer 180
QY 780 CATTTTCCATAC-----ATTAAGAGATGATCAT 806
Db 181 HisPheProTyrSerGlnTyrGlnPheTrpLysAsnPheGlnThrLeuLysIleValI 200
QY 807 CTGGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 866
Db 200 eLeuGlyLeuValLeuProLeuLeuValMetValIleCysTyrSerGlyIleLeuLysTh 220
QY 867 TCTGCTTGGTGTGCAATAG 926
Db 220 rLeuLeuArgCysArgAsnGlnLysLysArgHisArgAlaValArgLeuIlePheThrI 240
QY 927 CATGATGTTTATTTCTTCTGCGGCTCCCTACAAATGCTCTCTCTCTCTCTCTCTCTCT 986
Db 240 eMetIleValTyrPheLeuPheTrpAlaProTyrAsnIleValLeuLeuLeuAsnThrPh 260
QY 987 CCAGGAATCTTTGGCTGAATATTGACAGTCTTAACAGGTTCGACAGGTTCGACAGGTTC 1046
Db 260 eGlnGluPhePheGlyLeuAsnAsnCysSerSerSerAsnArgLeuAspGlnAlaMetG 280

RESULT 2

US-08-861-105-14
Sequence 14, Application US/08861105
Patent No. 6258527

GENERAL INFORMATION:

APPLICANT: LITTMAN, DAN R.
APPLICANT: DENG, HONGKUI
APPLICANT: ELLMEIER, WILFRIED
APPLICANT: LANDAU, NATHANIEL R.
APPLICANT: LIU, RONG

TITLE OF INVENTION: G-COUPLED RECEPTORS ASSOCIATED WITH
MACROPHAGE-TROPIC HIV, AND DIAGNOSTIC AND THERAPEUTIC
USES THEREOF
TITLE OF INVENTION: G-COUPLED RECEPTORS ASSOCIATED WITH
MACROPHAGE-TROPIC HIV, AND DIAGNOSTIC AND THERAPEUTIC
USES THEREOF

CORRESPONDENCE ADDRESS:

ADDRESSEE: David A. Jackson, Esq.
STREET: 411 Hackensack Ave, Continental Plaza, 4th
FLOOR
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/861,105

FILING DATE:

CLASSIFICATION: 436

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/666,020

FILING DATE: 19-JUN-1996

CLASSIFICATION: 436

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/227,319

FILING DATE: 13-APR-1994

ATTORNEY/AGENT INFORMATION:

NAME: Jackson Esq., David A.

REGISTRATION NUMBER: 26,742

REFERENCE/DOCKET NUMBER: 1049-1-004 N1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 201-487-5800

TELEFAX: 201-343-1684

INFORMATION FOR SEQ ID NO: 14:

SEQUENCE CHARACTERISTICS:

LENGTH: 352 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

ORGANISM: Homo sapiens

US-08-861-105-14

Alignment Scores:

Pred. No.: 5,68e-182 Length: 352
 Score: 1745.50 Matches: 340
 Percent Similarity: 96.60% Conservative: 1
 Best Local Similarity: 96.32% Mismatches: 0
 Query Match: 66.42% Indels: 12
 DB: 4 Gaps: 1

US-09-938-703-3 (1-1442) x US-08-861-105-14 (1-352)

```

QY 240 ATGGATTATCAAGTCTCAAGTCAATCTATGACATCAATATTATATACATCGGACCCCTGC 299
Db 1 MetAspTyrGlnValSerSerProIleTyrAspIleAsnTyrTyrThrSerGluProCys 20
QY 300 CAAAAATCAATGTGAAGCAATCGGACCGCCCTCCCTCGCTCCCTACTACTACACGGTG 359
Db 21 GlnIysIleAsnValIysGlnIleAlaAlaArgLeuLeuProLeuTyrSerLeuVal 40
QY 360 TTCATCTTTGGTTTGGGCAACATGCTGGTCACTCCCTCATCTGATTAATCACTGACAAAGG 419
Db 41 PheIlePheGlyPheValIysAsnMetLeuValIleLeuIleLeuIleAsnCysLysArg 60
QY 420 CTGAAGACATGACGACATCTACCTGCTCAACCTGCCATCTCTGACCTGTTTTCCTT 479
Db 61 LeuLysSerMetThrAspIleTyrLeuLeuAsnLeuAlaIleSerAspLeuPhePheLeu 80
QY 480 CTACTGTCCCTTCTGGGCTCACTATGCTGCCGCCAGTGGGACCTTTGGAAATPACAAAG 539
Db 81 LeuThrValProPheThrPalaHisTyrAlaAlaLagInTrpAspPheGlyAsnThrMet 100
QY 540 TGTCACTCTTGACAGGCTCTATTATTAGGCTCTTCTGGAATCTTCTTCATCATC 599
Db 101 CysGlnLeuLeuThrGlyLeuTyrPheIleGlyPheSerGlyIlePhePheIleIle 120
QY 600 CTCCTGACAAATCAGTAGTACCTGGCTGCTGCCATGCTGTGTTTCTTTAAAGCCAGG 659
Db 121 LeuLeuThrIleAspArgTyrLeuAlaValValHisAlaValPheAlaLeuLysAlaArg 140
QY 660 ACGTCACTTTGGGTTGGGACAAAGTGTATCATCTGGTGGTGGCTGTGTTGGCTCT 719
Db 141 ThrValThrPheGlyValValThrSerValIleThrTrpValValAlaValPheAlaSer 160
QY 720 CTCCTGAGATCATCTTTACAGATCTCAAAAGAGTCTTCAATACACCTGCAGTCT 779
Db 161 LeuProGlyIleIlePheThrArgSerGlnLysGlnGlyLeuHisTyrThrCysSerSer 180
QY 780 CATTTTCCATAC-----ATTAAAGATAGTCAT 806
Db 181 HisPheProTyrSerGlnTyrGlnPheTrpLysAsnPheGlnThrLeu-LysIleValI 200
QY 807 CTGGGGCTGGCTCGCGCTGCTGTGATGCTATGCTACTGCTGGAATCTTAAAC 866
Db 200 eLeuGlyLeuValLeuProLeuLeuValMetValIleCysTyrSerGlyIleLeuLysTh 220
QY 867 TCTGCTTCGGTTCGAAATGAGAGAGGACAGGCTGTGAGGCTTATCTTCAACAT 926
Db 220 rLeuLeuArgCysArgAsnGlnLysLysArgHisArgAlaValArgLeuIlePheThrII 240
QY 927 CATGATTTGTTATTTTCTCTCTGGCTCCCTACACATGCTCTTCTCCCTGACACCTT 986
Db 240 eMetIleValTyrPheLeuPheTrpAlaProTyrAsnIleValLeuLeuLeuAsnThrPh 260
QY 987 CCAGAAATCTTTGGCTCGAATATTGCAGTAGCTTAACAGGTTCGACCAAGTATGCA 1046
Db 260 eGlnGlnPhePheGlyLeuAsnAsnCysSerSerSerAsnArgLeuAspGlnAlaMetGI 280
QY 1047 GGTGACAGACTCTTGGGATGAGCCACTGCTGCATCAACCCCATCATCTATGCTTTGT 1106
Db 280 nValThrGluThrLeuGlyMetThrHisCysCysIleAsnProIleTyrAlaPheVa 300
QY 1107 CGGGGAGAACTCAGAACTACCTTCTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1166
Db 300 IglYgluLysPheArgAsnTyrLeuLeuValPhePheGlnLysHisIleAlaLysArgPh 320

```

```

QY 1167 CTCGAAATGCTGTCTATTTCACAGAGGCTCCGAGGAGCAGCAGTCACTGTTACAC 1226
Db 320 eCysLysCysCysSerIlePheGlnGlnAlaProGluArgAlaSerSerValTyrTh 340
QY 1227 CCGATCCACTGGGAGCAGCAAAATATCTGTGGGCTTG 1263
Db 340 rArgSerThrGlyGlnGlnIleSerValGlyLeu 352

```

RESULT 3

```

US-08-575-967A-2
; Sequence 2, Application US/08575967A
; Patent No. 6265184
; GENERAL INFORMATION:
; APPLICANT: Gray et al.
; TITLE OF INVENTION: Chemokine Receptor Materials and Methods
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 S. Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/575,967A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6265184and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 32918
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-485-1900
; TELEFAX: 206-485-1662
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 352 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: /- "88c amino acid sequence"
US-08-575-967A-2

```

Alignment Scores:
 Pred. No.: 5,68e-182 Length: 352
 Score: 1745.50 Matches: 340
 Percent Similarity: 96.60% Conservative: 1
 Best Local Similarity: 96.32% Mismatches: 0
 Query Match: 66.42% Indels: 12
 DB: 4 Gaps: 1

US-09-938-703-3 (1-1442) x US-08-575-967A-2 (1-352)

```

QY 240 ATGGATTATCAAGTCTCAAGTCAATCTATGACATCAATATTATATACATCGGACCCCTGC 299
Db 1 MetAspTyrGlnValSerSerProIleTyrAspIleAsnTyrTyrThrSerGluProCys 20
QY 300 CAAAAATCAATGTGAAGCAATCGGACCGCCCTCCCTCGCTCCCTACTACTACTGCTG 359
Db 21 GlnIysIleAsnValIysGlnIleAlaAlaArgLeuLeuProLeuTyrSerLeuVal 40
QY 360 TTCATCTTTGGTTTGGGCAACATGCTGGTCACTCCCTCATCTGATTAATCACTGACAAAGG 419
Db 41 PheIlePheGlyPheValGlyAsnMetLeuValIleLeuIleLeuIleAsnCysLysArg 60

```

```

QY 420 CTGAAGAGATGACTGACATCTACCTGCTCAACCTGGCCATCTCTGACCTGTTTTCCTT 479
Db 61 LeuLysSerMetThrAspIleTyrLeuLeuAsnLeuAlaIleSerAspLeuPheLeu 80
QY 480 CTTACTCTCCCTTCTGGCTCACTATGCTGCCGCGAGTGGAGCTTTGGAAATACAAATG 539
Db 81 LeuThrValProPheTrpAlaHisTyrAlaAlaGlnTrpAspPheGlyAsnThrMet 100
QY 540 TGTCAACTCTTGACAGGCTCTATTTATAGGCTTCTCTCTGGAATCTTCTTCATCATC 599
Db 101 CysGlnLeuLeuThrGlyLeuTyrPheIleGlyPhePheSerGlyIlePhePheIle 120
QY 600 CTCTCGAATCGATAGTACCTGGCTGCTGCTCCATGCTGTGTTCTTTAAAGCCAGG 659
Db 121 LeuLeuThrIleAspArgTyrLeuAlaValHisAlaValPheAlaLeuLysAlaArg 140
QY 660 ACGGTCACCTTTGGGTGGTGACAAAGTGTGATCACTGGGTGGTGTGTTGGCTCT 719
Db 141 ThrValThrPheGlyValValThrSerValIleThrTrpValValAlaValPheAlaSer 160
QY 720 CTCCAGGAATCATCTTACAGATCTCAAAAGAGGTCTTCATTCACCTGCAGCTCT 779
Db 161 LeuProGlyIleIlePheThrArgSerGlnLysGluGlyLeuHisTyrThrCysSer 180
QY 780 CATTTTCCATAC-----ATTAAAGATAGTCAT 806
Db 181 HisPheProTyrSerGlnTyrGlnPheTrpLysAsnPheGlnThrLeu-LysIleValII 200
QY 807 CTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 866
Db 200 eLeuGlyLeuValLeuProLeuLeuValMetValIleCysTyrSerGlyIleLeuLysTh 220
QY 867 TCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 926
Db 220 rLeuLeuArgCysArgGlnLysLysArgHisArgAlaValArgLeuIlePheThrII 240
QY 927 CATGATGTTTATTTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 986
Db 240 eMeIleValTyrPheLeuPheTrpAlaProTyrAsnIleValLeuLeuLeuAsnThrPh 260
QY 987 CAGGAATCTTTGGCTGGAATATGACAGTCTTACAGGTTGGACCAAGCTATGCA 1046
Db 260 eGlnGluPhePheGlyLeuAsnAsnCysSerSerAsnArgLeuAspGlnAlaMetGI 280
QY 1047 GTCACAGACTCTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 1106
Db 280 nValThrGluThrLeuGlyMetThrHisCysCysIleAsnProIleIleTyrAlaPheVa 300
QY 1107 CGGGGAGAGTTCCAGAACTACCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1166
Db 300 lGlyLysPheArgAsnTyrLeuLeuValPhePheGlnLysHisIleAlaLysArgPh 320
QY 1167 CTGCAATGCTGTTCTTATTTCCAGCAGAGGCTCCGAGCAGAGCTCAGTTTACAC 1226
Db 320 eCysLysCysCysSerIlePheGlnGlnAlaProGluArgAlaSerSerValTyrTh 340
QY 1227 CGATCCACTGGGAGGAGGAATATCTGTTGGGCTTG 1263
Db 340 rArgSerThrGlyGluGlnGluIleSerValGlyLeu 352

```

RESULT 4

```

US-08-833-752-5
; Sequence 5, Application US/08833752
; Patent No. 6448375
; GENERAL INFORMATION:
; APPLICANT: SAMSON, MICHEL
; APPLICANT: PARMENTIER, MARC
; APPLICANT: VASSART, GILBERT
; APPLICANT: LIBERT, FREDERICK
; TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR
; TITLE OF INVENTION: AND NUCLEIC ACID MOLECULES ENCODING SAID RECEPTOR
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:

```

```

; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/833,752
; FILING DATE: 9-APR-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Altman, Daniel E.
; REGISTRATION NUMBER: 34,115
; REFERENCE/DOCKET NUMBER:
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 352 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-833-752-5

```

```

Alignment Scores:
Pred. No.: 5 68e-182 Length: 352
Score: 1745.50 Matches: 340
Percent Similarity: 96.60% Conservative: 1
Best local Similarity: 96.32% Mismatches: 0
Query Match: 66.42% Indels: 12
DB: 4 Gaps: 1
US-09-938-703-3 (1-1442) x US-08-833-752-5 (1-352)

```

```

QY 240 ATGGATTATCAAGTGTCAAGTCCCAATCTATGACATCAATATTATATACATGAGGCGCTTC 299
Db 1 MetAspTyrGlnValSerSerProIleTyrAspIleAsnTyrThrSerGluProCys 20
QY 300 CAAAATAATCAATGTAAGCAAAATCGACCGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCT 359
Db 21 GlnLysIleAsnValLysGlnIleAlaAlaArgLeuProLeuTyrSerLeuVal 40
QY 360 TTTACTTTTGGTTTGGGCAACATGCTGCTCATCTCTCATCTGATAAACTGCAAAAGG 419
Db 41 PheIlePheGlyPheValGlyAsnMetLeuValIleLeuIleAsnCysLysArg 60
QY 420 CTGAAGACCATGACTGACATCTACCTGCTCAACCTGGCCATCTCTGACCTGTTTTCCTT 479
Db 61 LeuLysSerMetThrAspIleTyrLeuLeuAsnLeuAlaIleSerAspLeuPhePheLeu 80
QY 480 CTTACTGCTCCCTCTGGGCTCACTATGCTGCGCCGAGTGGAGCTTTGGAAATACAAATG 539
Db 81 LeuThrValProPheTrpAlaHisTyrAlaAlaGlnTrpAspPheGlyAsnThrMet 100
QY 540 TGTCAACTCTTTGACAGGCTCTATTTATAGGCTTCTCTCTGGAATCTTCTTCATCATC 599
Db 101 CysGlnLeuLeuThrGlyLeuTyrPheIleGlyPhePheSerGlyIlePhePheIle 120
QY 600 CTCTCGAATCGATAGTACCTGGCTGCTGCTCCATGCTGCTGCTGCTGCTGCTGCTGCT 659
Db 121 LeuLeuThrIleAspArgTyrLeuAlaValHisAlaValPheAlaLeuLysAlaArg 140
QY 660 ACGGTCACCTTTGGGTGGTGACAAAGTGTGATCACTGGGTGGTGTGTTGGCTCT 719
Db 141 ThrValThrPheGlyValValThrSerValIleThrTrpValValAlaValPheAlaSer 160
QY 720 CTCCAGGAATCATCTTACAGATCTCAAAAGAGGTCTTCATTCACCTGCAGCTCT 779
Db 161 LeuProGlyIleIlePheThrArgSerGlnLysGluGlyLeuHisTyrThrCysSer 180

```



```

Db      280  nValThrGluThrLeuGlyMetThrHisCysCysIleAsnProIleIleTyrAlaPheVa 300
QY      1107  CGGGGAGAAGTTCAGAACTACCTCTTACTCTTCTCCAAAGACACATTCGCAACGCTT 1166
Db      300  lGlyGluLysPheArgAsnTyrLeuLeuValPhePheGlnLysHisIleAlaLysArgPh 320
QY      1167  CTGCAAAAGCTGTTCATTTCACGAAGAGGCTCCGCGAGCAGCAAGCTCAGCTTTACAC 1226
Db      320  eCysLysCysSerIlePheGlnGlnAlaProGluArgAlaSerValTyrThr 340
QY      1227  CGATCCACTGGGGAGCAGGAATAATCTCTGGGCTTG 1263
Db      340  rArgSerThrGlyGlnGlnIleSerValGlyLeu 352

RESULT 6
US-09-534-185-52
; Sequence 52, Application US/09534185
; Patent No. 6403767
; GENERAL INFORMATION:
; APPLICANT: Graham, Gerard J. et al.
; TITLE OF INVENTION: No. 6403767el Molecules of the G Protein-Coupled
; Heptahelical Receptor Superfamily and Uses
;

```

Patent No. 6025154
 GENERAL INFORMATION:
 APPLICANT: LI, YI
 TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING HUMAN G-PROTEIN
 TITLE OF INVENTION: CHEMOKINE RECEPTOR HDGNR10 (AS AMENDED)
 NUMBER OF SEQUENCES: 9
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.
 STREET: 1100 NEW YORK AVE., NW, SUITE 600
 CITY: WASHINGTON
 STATE: DC
 COUNTRY: USA
 ZIP: 20005
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/466,343D
 FILING DATE: 08-JUN-1995
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: STEFFE, ERIC K.
 REGISTRATION NUMBER: 36,688
 REFERENCE/DOCKET NUMBER: 1488.1150000/EKS/KLM
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 371-2600
 TELEFAX: (202) 371-2540
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 352 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-466-343D-2

Alignment Scores:
 Pred. No.: 2,46e-180 Length: 352
 Score: 1730.50 Matches: 336
 Percent Similarity: 96.32% Conservative: 4
 Best Local Similarity: 95.18% Mismatches: 1
 Query Match: 65.85% Indels: 12
 DB: 3 Gaps: 1

US-09-938-703-3 (1-1442) x US-08-466-343D-2 (1-352)

QY 240 ATGATATCAAGTCTCAAGTCCAAATCTATGACATCAATATTATATACATCGAGCCCTGC 299
 Db 1 MetAspTyrGlnValSerSerProIleTyrAspIleAsnTyrTyrThrSerGluProCys 20
 QY 300 CAAAAATCAATGTAAGCAAAATCGCAGCCGCTCTCTCGCTCGCTACTCTACTGCTG 359
 Db 21 ProLysIleAsnValysGlnIleAlaAlaArgLeuLeuProLeuTyrSerLeuVal 40
 QY 360 TTCATCTTTGGTTTGGGCAACATCTGCTCATCTCTCTCATCTCTGATTAATCAAGG 419
 Db 41 PheIlePheGlyPheValGlyAsnMetLeuValIleLeuIleLeuIleAsnCysGlnArg 60
 QY 420 CTGAAGAGATGACATGACATCTACCTGCTCAACCTGGCCATCTCTGACCTGTTTCTCT 479
 Db 61 LeuGluSerMetThrAspIleTyrLeuLeuAsnLeuAlaIleSerAspLeuPheLeu 80
 QY 480 CTCTACTGCTCTCTGCTGCTCACTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 539
 Db 81 LeuThrValProPheTrpAlaHisTyrAlaAlaAlaGlnTrpAspPheGlyAsnThrMet 100
 QY 540 TGTCAACTCTTGACAGGCTCTATTATAGGCTTCTCTCTGGAATCTTCTTCATCATC 599
 Db 101 CysGlnLeuLeuThrGlyLeuTyrPheIleGlyPhePheSerGlyIlePhePheIle 120
 QY 600 CTCCTGCAATGATAGGTACCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 659

Db 121 LeuLeuThrIleAspArgTyrLeuAlaIleValHisAlaValPheAlaLeuLysAlaArg 140
 QY 660 AGGTCACCTTTGGGCTGGTGGTGAAGTGTGATCACTTGGGTGGTGGTGTGTGTGGCTCT 719
 Db 141 ThrValThrPheGlyValValThrSerValIleThrTrpValValAlaValPheAlaSer 160
 QY 720 CTCCTGAGATCATCTTACCAGATCTCAAAAGAGCTCTTCAATACACCTGAGCTCT 779
 Db 161 LeuProGlyIleIlePheThrArgSerGlnLysGluGlyLeuHisTyrThrCysSer 180
 QY 780 CATTTTCCATAC-----ATTAAGATAGTCA 806
 Db 181 HisPheProTyrSerGlnTyrGlnPheTrpLysAsnPheGlnThrLeu-LysIleVal 200
 QY 807 CTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 866
 Db 200 eLeuGlyLeuValLeuProLeuLeuValMetValIleCysTyrSerGlyIleLeuLys 220
 QY 867 TCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 926
 Db 220 rLeuLeuArgCysArgAsnGlnLysArgHisArgAlaValArgLeuIlePheThr 240
 QY 927 CATGATTTGTTTATTTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 986
 Db 240 eMetIleValTyrPheLeuPheTrpAlaProTyrAsnIleValLeuLeuAsnThrPh 260
 QY 987 CCAGGAATTTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1046
 Db 260 eGlnGluPhePheGlyLeuAsnAsnCysSerSerAsnArgLeuAspGlnAlaMetG 280
 QY 1047 GGTGACAGACTCTTGGGATGAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1106
 Db 280 nValThrGluThrLeuGlyMetThrHisCysCysIleAsnProIleIleTyrAlaPhe 300
 QY 1107 CCGGAGAGAGTTCAGAACTACCTCTTAGTCTTCTTCCAAAGACATTCGCAACGCTT 1166
 Db 300 lGlyLysPheArgAsnTyrLeuLeuValPhePheGlnLysHisIleAlaLysArgPh 320
 QY 1167 CTGCAATGCTGTTCTATTTCACAGAGAGCTCCCGAGGAGCAAGCTCAGTTTACAC 1226
 Db 320 eCysLysCysSerIlePheGlnGlnAlaProGluArgAlaSerValTyrTh 340
 QY 1227 CCGATCCACTGGGAGCAGGAAATATCTGTGGCTTG 1263
 Db 340 rArgSerThrGlyGluGlnGluIleSerValGlyLeu 352

RESULT 8
 US-09-517-605-5
 ; Sequence 5, Application US/09517605
 ; Patent No. 6391567
 ; GENERAL INFORMATION:
 ; APPLICANT: Littman, Dan R.
 ; APPLICANT: Kwon, Douglas S.
 ; APPLICANT: van Kooyk, Yvette
 ; APPLICANT: Geljtenbeck, Tseo
 ; TITLE OF INVENTION: METHODS OF USING A FACILITATOR OF RETROVIRAL ENTRY INTO
 ; TITLE OF INVENTION: CELLS
 ; FILE REFERENCE: 1049-1-017
 ; CURRENT APPLICATION NUMBER: US/09/517,605
 ; CURRENT FILING DATE: 2000-03-02
 ; NUMBER OF SEQ ID NOS: 17
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 5
 ; LENGTH: 352
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-517-605-5

Alignment Scores:
 Pred. No.: 5,02e-179 Length: 352
 Score: 1718.50 Matches: 334
 Percent Similarity: 95.75% Conservative: 4
 Best Local Similarity: 94.62% Mismatches: 3

Query Match: 65.39% Indels: 12
DB: 4 Gaps: 1
US-09-938-703-3 (1-1442) x US-09-517-605-5 (1-352)

QY 240 ATGGATTATCAAGTGTCAAGTCAATCTATGACATCAATATTATATACATCGAGCGCTTCG 299
Db 1 MetAspPheGlnValSerProThrTyrAspPheAspThrSerGluProCys 20
QY 300 CAAAAATCAATGTGAAGCAATCGCAGCGCGCTCTCTGCTCGCTCTACTCTACTCTGCTG 359
Db 21 GlnLysIleAsnValLysGlnIleAlaAlaArgLeuLeuProLeuTyrSerLeuVal 40
QY 360 TTTACTTTGGTTTGGGCAACATGCTGGTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 419
Db 41 PheIlePheGlyPheValGlyAsnMetLeuValIleLeuValLeuIleAsnLysArg 60
QY 420 CTGAAGAGCATGACATCTACCTGCTCAACCTGGCCATCTCTGACCTGTTTTCCTT 479
Db 61 LeuLysSerMetThrAspIleTyrLeuLeuAsnLeuAlaIleSerAspLeuPheLeu 80
QY 480 CTTACTGTCCTCTGGGCTCACTATGCTGCGCCCGAGTGGGACTTTGGAATACAATG 539
Db 81 LeuThrValProPheTrpAlaHisTyrAlaAlaGlnTrpAspPheGlyAsnThrMet 100
QY 540 TGTCAACTCTGTACAGGCTCTATTTATAGGCTTCTCTCTCTCTCTCTCTCTCTCTCT 599
Db 101 CysGlnLeuLeuThrGlyLeuTyrPheIleGlyPheSerGlyIlePhePheIle 120
QY 600 CTCTCGAATCATGATAGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 659
Db 121 LeuLeuThrIleAspArgTyrLeuAlaIleValHisAlaValPheAlaLeuLysAlaArg 140
QY 660 AGGTCACCTTTGGGCTGAGCAAGTGTGATCACTTGGTGGTGGCTGTGTGGTGTCT 719
Db 141 ThrValThrPheGlyValValThrSerValIleThrTrpValValAlaValPheAlaSer 160
QY 720 CTCCAGGAATCATCTTTACAGATCTCAAAAAGAGTCTTCTTACACCTGACGCTCT 779
Db 161 LeuProGlyIlePhePheThrArgSerGlnLysGluGlyLeuHisTyrThrCysSer 180
QY 780 CATTTTCATAC-----ATTAAAGATAGTCAT 806
Db 181 HisPheProTyrSerGlnTyrGlnPheTrpLysAsnPheGlnThrLeu-LysIleValI 200
QY 807 CTGGGCTGGTCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 866
Db 200 eLeuGlyLeuValLeuProLeuLeuValMetValIleCysTyrSerGlyIleLeuLysTh 220
QY 867 TCTGCTCGGTGTGAAATGAGAAGAGGACAGGCTGTGAGGCTTATCTTCACCAT 926
Db 220 rLeuLeuArgCysArgAsnGlnLysLysArgHisArgAlaValArgLeuIlePheThrI 240
QY 927 CATGATGTTTATTCTCTCTGGGCTCCCTACAACTGTGCTCTCTCTCTCTCTCTCTCTCT 986
Db 240 eMetIleValTyrPheLeuPheTrpAlaProTyrAsnIleValLeuLeuAsnThrPh 260
QY 987 CCAGGAATCTTTGGCTGAATAATTCAGTAGCTCAACAGTGTGACCAAGCTATGCA 1046
Db 260 eGlnGluPhePheGlyLeuAsnGlnCysSerSerAsnArgLeuAspGlnAlaMetG 280
QY 1047 GGTGACAGACTTTGGGATGAGCAGCTGCTGATCAACCCCATCATCTATGCTCTTGT 1106
Db 280 nValThrGluThrLeuGlyMetThrHisCysCysIleAsnProIleIleTyrAlaPheVa 300
QY 1107 CGGAGAGAACTCAGAACTACCTCTTATGCTCTTCTTCCAAAGCAACATGCGCAACGCTT 1166
Db 300 lGlyGluLysPheArgAsnTyrLeuLeuValPhePheGlnLysHisIleAlaLysHisPh 320
QY 1167 CTGCAAAATGCTTCTATTTTCCAGCAAGAGCTCCCGAGCGGAGCTCAGTTTACAC 1226
Db 320 eCysLysCysSerIlePheGlnGlnGluAlaProGluArgAlaSerSerValTyrTh 340

QY 1227 CCGATCCATGGGAGCAGGAAATATCTGTGGCTTG 1263
Db 340 rArgSerThrGlyGluGlnGluIleSerValGlyLeu 352

RESULT 9
US-08-724-984A-2
; Sequence 2, Application US/08724984A
; Patent No. 6388055
; GENERAL INFORMATION:
; APPLICANT: Derk Bergsma, Mary Brawner, and Usman Shabon
; TITLE OF INVENTION: No. 6388055el Mouse Genomic Clone of the CC-
; TITLE OF INVENTION: CARS Receptor
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road, P.O. Box 1539
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
; COMPUTER: IBM 486
; OPERATING SYSTEM: WINDOWS FOR WORKGROUPS
; SOFTWARE: MICROSOFT WORD
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/724.984A
; FILING DATE: October 3, 1996
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: William T. Han
; REGISTRATION NUMBER: 34,344
; REFERENCE/DOCKET NUMBER: ATG50023
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610 270 5024
; TELEFAX: 610 270 5090
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 354
; TYPE: Amino Acid
; TOPOLOGY: Linear
; US-08-724-984A-2

Alignment Scores:
Pred. No.: 3,42e-151 Length: 354
Score: 1463.50 Matches: 279
Percent Similarity: 86.48% Conservative: 28
Best Local Similarity: 78.59% Mismatches: 34
Query Match: 55.69% Indels: 14
DB: 4 Gaps: 2

US-09-938-703-3 (1-1442) x US-08-724-984A-2 (1-354)

QY 240 ATGGATTATCAAGTGTCAAGTCAATCTATGACATCAATATTATATACATCGAG 293
Db 1 MetAspPheGlnGlySerValProThrTyrIleTyrAspPheAspThrGlyMetSerAla 20
QY 294 CCTGCAAAAAATCAATGTGAAGCAATCGCAGCGCGCTCTCTGCTCGCTCTACTCA 353
Db 21 ProCysGlnLysIleAsnValLysGlnIleAlaGlnLeuLeuProLeuTyrSer 40
QY 354 CTGTTGTTCACTCTGTTGTTTGTGGCAACATGCTGCTCATCTCTCTCTCTCTCTCTCTCT 413
Db 41 LeuValPheIlePheGlyPheAlaGlyAsnMetMetValPheLeuIleLeuIleSerCys 60
QY 414 AAAGGCTGAAGAGCATGACTGACATCTACTCTCAACCTGGCCATCTCTGACCTGTTT 473
Db 61 LysLysLeuLysSerValThrAspIleTyrLeuAsnLeuAlaIleSerAspLeuLeu 80
QY 474 TTCCCTTCTACTGTCCCTTCTTGGGCTCACTATGCTGCGCCGAGTGGGACTTTGMAAT 533

```
Db 81 PheLeuLeuThrLeuProPheThrPalahisIyrAlaAlaSnGluTrpIlePheGlyAsn 100
QY 534 ACAATGTGCAACTTCAGACAGGCTCTATTTATAGGCTTCTCTCTGGAATCTCTTC 593
Db 101 IleMetCysLysValPheThrGlyValThrHisIleGlyTyrPheGlyGlyIlePhePhe 120
QY 594 ATCATCTCTGCAACATGATAGTACTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 553
Db 121 IleIleLeuLeuThrIleAspArgTyrIleAlaIleValHisAlaValPheAlaLeuLys 140
QY 654 GCCAGACGGTCACTCTGGGTGGTGCAGAGTGTGATGATGCTGGTGGTGGTGGTGT 713
Db 141 ValThrThrValAsnPheGlyValIleThrSerValValThrIrpValAlaValPhe 160
QY 714 GCGTCTCTCCAGCAATCATCTTTACAGATCTCAAAAGAGGCTTCATTCACCTGC 773
Db 161 AlaserLeuProGluIleIlePheThrArgSerGlnLysGluGlyPheHisTyrThrCys 180
QY 774 AGCTCTCATTTTCATAC-----ATTAAGAT 800
Db 181 SerProHisPheProHisThrGlnTyrHisPheTyrPheThrLeuLysLeu 200
QY 801 AGTCATCTTTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 860
Db 200 tValIleLeuSerLeuLeuLeuProLeuLeuValMetIleIleCysTyrSerGlyIle 220
QY 861 AAAAATCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 920
Db 220 uHisThrLeuPheArgCysArgAsnGluLysArgHisArgAlaValArgLeuIlePhe 240
QY 921 CACATCATGATTTTATTTCTCTCTGGGCTCCCTACACATGTCCTCTCTCTCTGAA 980
Db 240 eAlaIleMetIleValIyrPheLeuPheThrThrProTyrAsnIleValLeuLeuLeu 260
QY 981 CACCTTCCAGGAATCTTTGGCTGAATAATGTCAGTGTCTCAACAGTTGGACCAAGC 1040
Db 260 rThrPheGlnGluPhePheGlyLeuAsnAspCysSerSerSerAsnArgLeuAspGlnAl 280
QY 1041 TATGAGGTCAGACAGACTCTGGGATGACGACTGTCTGCTGCTGCTGCTGCTGCTGCT 1100
Db 280 aMetGlnAlaThrGluThrLeuGlyMetThrHisCysCysLeuAsnProValIleTyrAl 300
QY 1101 CTCTGCTGGGAGGTTGAGAACTACTCTTGTAGTCTCTTCCAAAGACACATTCGCA 1160
Db 300 aPheValGlyGluLysPheArgSerTyrLeuSerValPhePheArgHisIleVally 320
QY 1161 ACGCTTCTGCAATGCTGTTCTATTTCCAGCAAGAGGCTCCGAGCGAGCAAGCTCAGT 1220
Db 320 sArgPheCysLysArgCysSerIlePheGlnAspAsnProAspArgValSerVa 340
QY 1221 TTACACCGATCCACTGGGAGCAGGAATAATCTGTGGGCTGG 1263
Db 340 lTyrThrArgSerThrGlyGluHisGluValSerThrGlyLeu 354
```

RESULT 10

US-08-461-244-3

; Sequence 3, Application US/08461244

; Patent No. 5776729

; GENERAL INFORMATION:

; APPLICANT: Soppet, Daniel R.

; APPLICANT: Yi, Li

; APPLICANT: Ruben, Steven M.

; APPLICANT: Rosen, Craig A.

; TITLE OF INVENTION: HUMAN G-PROTEIN RECEPTOR HGBR32

; NUMBER OF SEQUENCES: 7

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN, CECCHI,

; ADDRESS: STUART & OLSTEIN

; STREET: 6 Becker Farm Road

; CITY: Roseland

; STATE: New Jersey

; COUNTRY: USA

```
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,244
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferraro, Gregory D.
; REGISTRATION NUMBER: 36,134
; REFERENCE/DOCKET NUMBER: 325800-445
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 347 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-461-244-3

Alignment Scores:
Pred. No.: 9,81e-137 Length: 347
Score: 1331.00 Matches: 258
Percent Similarity: 83.76% Conservative: 36
Best Local Similarity: 73.50% Mismatches: 46
Query Match: 50.65% Indels: 11
DB: 1 Gaps: 3

US-09-938-703-3 (1-1442) x US-08-461-244-3 (1-347)
```

```
QY 234 AACAGATGATTATCAAGTGTCAAGTCCAACTCATGACATCAATATTATATACATCGAG 293
Db 1 AsnGluSerGlyGluGluValThrThr---PhePheAspTyrAspTyr-----GlyAla 17
QY 294 CCCTGCCCAAAATCAATGTGAAGCAAAATCGGAGCCGCTCTCTCTCTCTCTCTCTCT 353
Db 18 ProCysHisLysPheAspValLysGlnIleGlyAlaGlnLeuLeuProLeuTyrSer 37
QY 354 CTGGTGTCTATCTCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 413
Db 38 LeuValPheIlePheGlyPheValGlyAsnMetLeuValLeuIleLeuLeuLeuLeu 57
QY 414 AAAAGGCTGAAGAGCATGACATCTACCTGCTCAACCTGCCATCTCTGCTGCTGCTTT 473
Db 58 LysLysLeuLysCysLeuThrAspIleTyrLeuLeuAsnLeuAlaIleSerAspLeuLeu 77
QY 474 TTCCTTCTTACTGCTCCCTCTCTGGGCTCTACTATGCTGCCGCCAGTGGGACCTTTGGA 533
Db 78 PheLeuIleThrLeuProLeuTyrAlaHisSerAlaAlaAsnGluTrpValPheGlyAsn 97
QY 534 ACAATGTGCAACTCTTGACAGGCTCTATTTATAGGCTTCTCTCTCTCTCTCTCTCT 593
Db 98 AlaMetCysLysLeuPheThrGlyLeuTyrHisIleGlyTyrPheGlyGlyIlePhePhe 117
QY 594 ATCATCTCTCTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 653
Db 118 IleIleLeuLeuThrIleAspArgTyrLeuAlaIleValHisAlaValPheAlaLeuLys 137
QY 654 GCCAGGAGGTCACCTTTGGGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 713
Db 138 AlaArgThrValThrPheGlyValValThrSerValIleThrIrpLeuValAlaValPhe 157
QY 714 GCGTCTCTCCAGCAATCATCTTTACAGATCTCAAAAGAGGCTTCATTCACCTGC 773
Db 158 AlaserValProGlyIleIlePheThrThrLysCysGlnLysGluAspSerValTyrValCys 177
QY 774 AGC-----TCTCATTTTCCATATCAATTAAGATAGTATCTTTGGG 812
```

| | | |
|------|--|------|
| 178 | GlyProTyrPheProArgGlyTrpAsnAsn-PheHisThrIleMetArgAsnIleLeuG1 | 137 |
| 813 | GCTGTCTCTCGCGTCTGTTCATGGTCATCTGTCTACGGGAATCCTTAAAACTCTGCT | 872 |
| 197 | YleuValLeuProLeuIleuMetValIleCysTyrSerGlyIleLeuIlysThrLeuIe | 217 |
| 873 | TCGGTGTCCGAATGAGAAAGAGGACAGGCGGTGAGCGTATCTTCCACCATCATGAT | 932 |
| 217 | uArgCysArgAsnGluIysArgHisArgAlaValArgValIlePheThrIleMetI1 | 237 |
| 933 | TGTTTATTTCTCTCTGGGTCCTCTACAACTTGTCTCTCTCTGGAACCTTCCAGGA | 992 |
| 237 | eValTyrPheLeuPheTrpThrProTyrAsnIleValIleLeuLeuAsnThrPheGlnG1 | 257 |
| 993 | ATCTTTTGGGCTGAATATTTCAGTACTCTAACAGGTGGACCAAGCTATGCAGGTGAC | 1052 |
| 257 | uPhePheGlyLeuSerAsnCysGluSerThrSerGlnLeuAspGlnAlaThrGlnValTh | 277 |
| 1053 | AGAGACTCTTGGGATGAGCGACTCTGTCATCAAGCCCATCTCTATGCGCTTCTGGGGA | 1112 |
| 277 | rGluThrLeuGlyMetThrHisCysCysIleAsnProIleIleTyrAlaPheValGlyG1 | 297 |
| 1113 | GAAGTTCAGAAACTACTCTTAGTCTCTTCTCCAAAGCACATTCGCAAGCTCTTGCA | 1172 |
| 297 | uLysPheArgTyrIleuSerValPhePheArgLysHisIleThrLysArgPheCysLys1 | 317 |
| 1173 | ATGCTGTCTATTTTCCAGCAAGAGGCTCCGAGGGAGCAAGCTCAGTTTACACCGATC | 1232 |
| 317 | sGlnCysProValPheTyrArgGluThrValAspGlyValThrSerThrAsnThrProSe | 337 |
| 1233 | CAC TGGGAGCAGGAATATCTGTGGCGTTG | 1263 |
| 337 | rThrGlyGluGlnGluValSerAlaGlyLeu | 347 |

RESULT 11

| | |
|--|--|
| US-08-450-393A-4 | |
| ; Sequence 4, Application US/08450393A | |
| ; Patent No. 5707815 | |
| ; GENERAL INFORMATION: | |
| ; APPLICANT: Charo, Israel | |
| ; APPLICANT: Coughlin, Shaun | |
| ; TITLE OF INVENTION: MAMMALIAN MONOCYTE CHEMOATTRACTANT | |
| ; TITLE OF INVENTION: PROTEIN RECEPTORS | |
| ; NUMBER OF SEQUENCES: 14 | |
| ; CORRESPONDENCE ADDRESS: | |
| ; ADDRESSEE: Cooley Godward Castro Huddleson & Tatum | |
| ; STREET: 5 Palo Alto Square | |
| ; CITY: Palo Alto | |
| ; STATE: California | |
| ; COUNTRY: USA | |
| ; ZIP: 94306-2155 | |
| ; COMPUTER READABLE FORM: | |
| ; MEDIUM TYPE: Floppy disk | |
| ; COMPUTER: IBM PC compatible | |
| ; OPERATING SYSTEM: PC-DOS/MS-DOS | |
| ; SOFTWARE: Patentin Release #1.0, Version #1.25 | |
| ; CURRENT APPLICATION DATA: | |
| ; APPLICATION NUMBER: US/08/450,393A | |
| ; FILING DATE: May 25, 1995 | |
| ; CLASSIFICATION: A24 | |
| ; ATTORNEY/AGENT INFORMATION: | |
| ; NAME: Cserr, Ioann | |
| ; REGISTRATION NUMBER: 31,822 | |
| ; REFERENCE/DOCKET NUMBER: UCAL-237/020S | |
| ; TELECOMMUNICATION INFORMATION: | |
| ; TELEPHONE: 415-843-5165 | |
| ; TELEFAX: 415-8857-0663 | |
| ; TELEX: 380816CooleyPA | |
| ; INFORMATION FOR SEQ ID NO: 4: | |
| ; SEQUENCE CHARACTERISTICS: | |
| ; LENGTH: 360 amino acids | |
| ; TYPE: amino acid | |

| | | | |
|-----------|------|--|------|
| <i>QY</i> | 1113 | GAGTTCAGAAACTACCTCTTAGTCTCTTCCAAAGCACATGCGCAACGCTTCGGAA | 1172 |
| | | :: | |
| <i>Db</i> | 310 | uIysPheArgTyLeuSerValPheArgLyHisIIleThrLysArgPheCysLy | 330 |
| | | | |
| <i>QY</i> | 1173 | ATGCGTGTATTATTTCCAGCAAGAGGCTCCCGAGGACAAGCTCAGTTTACACCGCATC | 1232 |
| | | | |
| <i>Db</i> | 330 | sGLncysProvalPheTyArgGluthrValAspGlyValThrSerThrAnThrProSe | 350 |
| | | | |
| <i>QY</i> | 1233 | CAC TGGGAGCAGGAATAATCTGTGGCGTG | 1263 |
| | | | |
| <i>Db</i> | 350 | rThrGlyGlnGlnIaValSerAlaGlyLeu | 360 |
| | | | |

RESULT 12

```

US-08-446-669-4
/ Sequence 4, Application US/08446669
/ Patent No. 6132987
/ GENERAL INFORMATION:
/ APPLICANT: Charo, Israel
/ APPLICANT: Coughlin, Shaun
/ TITLE OF INVENTION: MAMMALIAN MONOCYTE CHEMOATTRACTANT
/ TITLE OF INVENTION: MAMMALIAN MONOCYTE CHEMOATTRACTANT
/ TITLE OF INVENTION: MAMMALIAN MONOCYTE CHEMOATTRACTANT
/ TITLE OF INVENTION: MAMMALIAN MONOCYTE CHEMOATTRACTANT
/ TITLE OF INVENTION: MAMMALIAN MONOCYTE CHEMOATTRACTANT
/ NUMBER OF SEQUENCES: 14
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Cooley Godward Castro Huddleson & Tatum
/ STREET: 5 Palo Alto Square
/ CITY: Palo Alto
/ STATE: California
/ COUNTRY: USA
/ ZIP: 94306-2155
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/446,669
/ FILING DATE: May 25, 1995
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Neeley, Richard
/ REGISTRATION NUMBER: 30,092
/ REFERENCE/DOCKET NUMBER: UCAL-237/01US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 415-843-5000
/ TELEFAX: 415-857-0663
/ TELEX: 380816Coolvpa
/ INFORMATION FOR SEQ ID NO. 4:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 360 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
US-08-446-669-4

```

US-09-938-703-3 (1-1442) x US-08-446-669-4 (1-360)

[illegible]

| | | | |
|------|----|--|------|
| 354 | QY | CTGGTGTTCATCTTGGTTTTGTGGGCACACATCGTCGTGCATCTCCTCATGAATAACATGC | 413 |
| 51 | Db | LeuValPheIlePheGlyPheValGlyAsnMetLeuValLeuLeuLeuLeuLeuAsnCys | 70 |
| 414 | QY | AAAGSGTGAAGAGCATGACTGACATCTACCTGCTGCATCAACCTGCCATCTCTGACCTGTTT | 473 |
| 71 | Db | LysLysLeuLysCysLeuThrAspIleTyrLeuLeuAsnLeuAlaIleSerAspLeuLeu | 90 |
| 474 | QY | TTTCCTTCTACTGTCCTCCCTTGTGGCTCTACTATGCTGCCCGCAGTGGAGCTTTGGAAT | 533 |
| 91 | Db | PheLeuIleThrLeuProLeuTyrAlaHisSerAlaAsnGluTyrPheGlyAsn | 110 |
| 534 | QY | ACAAATGTCACACTCTGTACAGGGCTCTATTATTAAGGCTTCTCTCTGGAACTCTTC | 593 |
| 111 | Db | AlaMetCysLysLeuPheThrGlyLeuTyrHisIleGlyTyrPheGlyGlyIlePhePhe | 130 |
| 594 | QY | ATCATCTCCGTCACAATCGATAGGTACCTGGCTGCTGCATCGTCATGCTGTGTGCTTTAA | 653 |
| 131 | Db | IleIleLeuLeuThrIleAspTyrLeuAlaIleValHisAlaValPheAlaLeuLys | 150 |
| 654 | QY | GCAGAGCGGTCACTTTGGGGTGGTCAAAAGTGTGATCATCTTGGTGGTGGCTGTGTT | 713 |
| 151 | Db | AlaArgThrValThrPheGlyValValThrSerValIleThrTrpLeuValAlaValPhe | 170 |
| 714 | QY | GGGTCTCTCCAGGATCATCTTTACAGATCTCAAAAGAGGCTCTCATATCACCTGC | 773 |
| 171 | Db | AlaSerValProGlyIleIlePheThrLysCysGlnLysGluAspSerValTyrValCys | 190 |
| 774 | QY | AGC-----TCTCATTTTCCATACATTAAAGTAGTCATCTTGGG | 812 |
| 191 | Db | GlyProTyrPheProArgGlyTrpAsnAsn-PheHisThrIleMetArgAsnIleLeuG | 210 |
| 813 | QY | GCTGTGCTCGCGCTGCTTGTGCATGGTCATCTGCTACTTCGGGAATCCTAAAAATCTGCT | 872 |
| 210 | Db | YleuValleuProleuLeuIleMetValIleCysTyrSerGlyIleLeuLysThrLeuLe | 230 |
| 873 | QY | TCGGTGTGGAATGAGAGAGAGCGACAGGGCTGTGAGGCTTATCTTCCACCAFCATG | 932 |
| 230 | Db | uATGcysArgAsnLulysLysArgHisArgAlaValArgValIlePheThrIleMetI | 250 |
| 933 | QY | TGTTTATTTTCTCTCTGGGCTCCCTACACATATGCTCTCTCTCTGCAACACCTTCAGGA | 992 |
| 250 | Db | eValTyrPheLeuPheThrThrProTyrAsnIleValIleLeuLeuAsnThrPheGlnG | 270 |
| 993 | QY | ATTCTTTTGGCTGAATAATTGCAGTACTCTAACAGGTTGGACCAAGCTATGCAGGTGAC | 1052 |
| 270 | Db | uPhePheGlyLeuSerAsnCysGluSerThrSerGlnLeuAspIleAlaThrGlnValTh | 290 |
| 1053 | QY | AGAGACTCTGGGATGACCGACTGCTGCATCAACCCCATCATATPAGCCTTTGTGCGEGA | 1112 |
| 290 | Db | rgLuthrLeuGlyMetThrHisCysCysIleAsnProIleIleTyrAlaPheValGlyG | 310 |
| 1113 | QY | GAAATTCAGAACTACTCTTAGTCTTCTTCCAAAGACATTTGCCAACGCTCTCTGAA | 1172 |
| 310 | Db | uLysPheArgArgTyrIleuSerValPhePheArgLysHisIleThrLysArgPheCysL | 330 |
| 1173 | QY | ATGCTCTTCTATTTTCCAGAGAGGCTCCCGAGCGAGCAAGCTCAGTTTACACCGCATC | 1232 |
| 330 | Db | sGlnCysProValPheTyrArgGluThrValAspGlyValThrSerThrAsnThrProse | 350 |
| 1233 | QY | CATCGGGAGACGGAATATCTGTGGGGCTTG | 1263 |
| 350 | Db | rThrGlyGluGlnGluValSerAlaGlyLeu | 360 |

RESULT 13

```

RESUM 13
US-09-045-583-50
; Sequence 50, Application US/09045583
; Patent No. 6287805
; GENERAL INFORMATION:
; APPLICANT: Graham, Gerard J. et al.
; TITLE OF INVENTION: No. 6287805el Molecules of the G Protein-Coupled
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:

```

ADDRESSEE: LAHIVE & COCKFIELD, LLP
 STREET: 28 State Street
 CITY: Boston
 STATE: Massachusetts
 COUNTRY: USA
 ZIP: 02109
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/045,583
 FILING DATE: 20-MAR-98
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Mandragoras, Amy E.
 REGISTRATION NUMBER: 36,207
 REFERENCE/POCKET NUMBER: MNI-044
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617)227-7400
 TELEFAX: (617)742-4214
 INFORMATION FOR SEQ ID NO: 50:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 360 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 FRAGMENT TYPE: internal
 US-09-045-583-50

Alignment Scores:

Pred. No.: 9.96e-137 Length: 360
 Score: 1331.00 Matches: 258
 Percent Similarity: 83.76% Conservative: 36
 Best Local Similarity: 73.50% Mismatches: 46
 Query Match: 50.65% Indels: 11
 DB: 4 Gaps: 3

US-09-938-703-3 (1-1442) x US-09-045-583-50 (1-360)

QY 234 AACAGATGATTTATCAAGTCTCAAGTCCAAATCTATGACATCAATATTATACATCGAG 293
 DB 14 AAGGluSerGlyGluGluValThr---PhePheAspTyrAspTyr-----GlyAla 30
 QY 294 CCCTGCCAAAATCAATGTGAAGCAATCGCAGCCCGCTCTCTCGCTCCGCTCTACTCA 353
 DB 31 ProcysHisLysPheAspValLysGlnIleGlyAlaGlnLeuLeuProLeuTyrSer 50
 QY 354 CTGGTGTCTATCTTTGGTTTGGGCAACATGCTGGTCATCTCTCATCTGATTAAGTGC 413
 DB 51 LeuValPheIlePheGlyPheValGlyAsnMetLeuValValLeuLeuLeuLeuAsnCys 70
 QY 414 AAAAGGCTGAAGACATGACATCTACCTGCTCAACCTGCGCATCTCTGACCTGTTT 473
 DB 71 LysLysLeuLysCysLeuThrAspIleTyrLeuLeuAsnLeuAlaIleSerAspLeuLeu 90
 QY 474 TTCCTTTTACTGCTCCCTCTCTGGGCTCACTATGTCGCCGCCAGCGGACCTTTGGAAT 533
 DB 91 PheLeuIleThrLeuProLeuTyrAlaHisSerAlaAlaAsnGluTrpValPheGlyAsn 110
 QY 534 ACATGTGTCACTCTTACAGGCTCTATTATTAGCTCTCTCTCTGGAATCTTCTTC 593
 DB 111 AlaMetCysLeuLeuPheThrGlyLeuTyrHisIleGlyTyrPheGlyIlePhePhe 130
 QY 594 ATCATCTCTCTGACATGATGATGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTTAA 653
 DB 131 IleIleLeuLeuThrIleAspTyrLeuAlaIleValHisAlaValPheAlaLeuLys 150
 QY 654 GCCAGGACGCTACCTTTTGGGTTGGTGACAAAGTGTATCACTTGGTGGTGGCTGTGTTT 713

DB 151 AlaArgThrValThrPheGlyValValThrSerValThrTrpLeuValAlaValPhe 170
 QY 714 GGTCTCTCCAGGAATCATCTTTACAGATCTCAAAAGAGGCTCTTATTACACCTGC 773
 DB 171 AlaSerValProGlyIleIlePheThrLysCysGlnLysGluAspSerValTyrValCys 190
 QY 774 AGC-----TCTCATTTTCCATACATTAAGATAGTATCTCTGGG 812
 DB 191 GlyProTyrPheProArgGlyTrpAsnAsn-PheHisThrIleMetArgAsnIleLeuG 210
 QY 813 GTGTCTCTCCGCTGCTGTCTGCTATCTCTACTCTGGGGAATCTCTAAACCTCTGCT 872
 DB 210 yLeuValLeuProLeuLeuIleMetValIleCysTyrSerGlyIleLeuLysThrIle 230
 QY 873 TCGGTGTGAAATGAGAAGAGGACAGGCTGTGAGGCTTATCTTACCATCATCAT 932
 DB 230 uArgCysArgAsnGluLysArgHisArgAlaValArgValIlePheThrIleMetI 250
 QY 933 TGTATTATTTCTCTCTGGGCTCCCTACAAATGTCCTCTCTCTGACACCTTCCAGGA 992
 DB 250 eValTyrPheLeuPheThrProTyrAsnIleValIleLeuLeuAsnThrPheSlnG 270
 QY 993 ATCTCTTGGCTGAAATGTCAGTCTTAACAGTGTGACAGTGTGACCAAGCTATGACAGTGC 1052
 DB 270 uPhePheGlyLeuSerAsnCysGluSerThrSerGlnLeuAspGlnAlaThrGlnVal 290
 QY 1053 AGAGACTCTTGGGATGACGCTGCTGATCAACCCATCATCTATGCTTGTCTGGGGA 1112
 DB 290 rGluThrLeuGlyMetThrHisCysCysIleAsnProIleIleTyrAlaPheValGlyG 310
 QY 1113 GAAGTTCAGAACTACCTCTTACTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1172
 DB 310 uLysPheArgTyrLeuSerValPhePheArgLysHisIleThrLysArgPheCysLy 330
 QY 1173 ATCTGCTTCTTATTTTCCAGCAAGGCTCCCGCAGCAGCAGCTCACTTTTACACCGGATC 1232
 DB 330 sGlnCysProValPheTyrArgGluThrValAspGlyValThrSerThrAsnThrPro 350
 QY 1233 CACTGGGAGCAGGAATATCTGTGGGCTG 1263
 DB 350 rThrGlyGluGlnGluValSerAlaGlyLeu 360

RESULT 14

US-09-534-185-50
 ; Sequence 50, Application US/09534185
 ; Patent No. 6403767
 ; GENERAL INFORMATION:
 ; APPLICANT: Graham, Gerard J. et al.
 ; TITLE OF INVENTION: No. 6403767el Molecules of the G Protein-Coupled
 ; Heptahelical Receptor Superfamily and Uses
 ; Therefor
 ;
 ; NUMBER OF SEQUENCES: 56
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: LAHIVE & COCKFIELD, LLP
 ; STREET: 28 State Street
 ; CITY: Boston
 ; STATE: Massachusetts
 ; COUNTRY: USA
 ; ZIP: 02109
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/534,185
 ; FILING DATE: 24-Mar-2000
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 09/045,583
 ; FILING DATE: <Unknown>
 ; ATTORNEY/AGENT INFORMATION:

NAME: Mandragoras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: MNI-044
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 50:
SEQUENCE CHARACTERISTICS:
LENGTH: 360 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
SEQUENCE DESCRIPTION: SEQ ID NO: 50:
US-09-534-185-50

Alignment Scores:
Pred. No.: 9,96e-137 Length: 360
Score: 1331.00 Matches: 258
Percent Similarity: 83.76% Conservative: 36
Best Local Similarity: 73.50% Mismatches: 46
Query Match: 50.65% Indels: 11
DB: 4 Gaps: 3

US-09-938-703-3 (1-1442) x US-09-534-185-50 (1-360)

QY 234 AACAGATGGATATCAAGTGTCAAGTCCAAATCTATGACATCAATATATATACATCGGAG 293
Db 14 AsnGluSerGlyGluGluValThrThr---PhePheAspTyrAspTyr-----GlyAla 30
QY 294 CCCGCGCAAAATCAATGTAAGCAAAATCGCAGCGCGCTCTGCTCGCTCCGCTACTCA 353
Db 31 ProCysHisLysPheAsnValLysGlnIleGlyAlaGlnLeuLeuProLeuTyrSer 50
QY 354 CTGGTGTTCATCTTTGGTTTGGCGCAATGCTGCTCATCTCTCATCTCTGATAAATGC 413
Db 51 LeuValPheIlePheGlyPheValGlyAsnMetLeuValValLeuIleLeuIleAsnCys 70
QY 414 AAAAGGCTGAAGAGCATGACATCTACCTGCTCAACCTGCCATCTCTGACCTCTTT 473
Db 71 LysLysLeuLysCysLeuThrAspIleTyrLeuLeuAsnLeuAlaIleSerAspLeuLeu 90
QY 474 TTCCTCTTACTGCCCTCTCTGGCTCACTATGCTGCCCGCAGTGGGACTTTGGAAT 533
Db 91 PheLeuIleThrLeuProLeuTyrAlaHisSerAlaAlaAsnGluTyrValPheGlyAsn 110
QY 534 ACAATGTGCACTCTTGACAGGGCTCTATTTATAGCTTCTCTCTGGAATCTTCTTC 593
Db 111 AlaMetCysLysLeuPheThrGlyLeuTyrHisIleGlyTyrPheGlyIlePhePhe 130
QY 594 ATCATCTCTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 653
Db 131 IleLeuLeuThrIleAspArgTyrLeuAlaIleValHisAlaValPheAlaLeuLys 150
QY 654 GCCAGGACGCTCACCTTTGGGGTGGTGCACAGTGTGATCATCTTGGTGGTGGTGGT 713
Db 151 AlaArgThrValThrPheGlyValValThrSerValIleThrTrpLeuValAlaValPhe 170
QY 714 GCGTCTCTCCAGGAATCACTTTACAGATCTCAAAAGAGGCTTTCATACACCTGC 773
Db 171 AlaSerValProGlyIleIlePheThrLysCysGlnLysGlnAspSerValTyrValCys 190
QY 774 AGC-----TCTCATTTTCCATACATTAAGATGATCACTGGG 812
Db 191 GlyProTyrPheProArgGlyTyrPAsnAsn-PheHisThrIleMetArgAsnIleLeuG 210
QY 813 GCTGCTCTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 872
Db 210 yLeuValLeuProLeuLeuIleMetValIleCysTyrSerGlyIleLeuLysThrLeuLe 230
QY 873 TCGGTGTGCAATGAGAAGAGGACAGGCGCTGAGGCTTATCTTCCACCATCATGAT 932
Db 230 uArgCysArgAsnGluLysLysArgHisArgAlaValArgValIlePheThrIleMetI 250

QY 933 TGTATTATTTCTTCTTGGGTCCCTACAAATATGCTCTTCTCTGACACCTTCCAGGA 992
Db 250 eValTyrPheLeuPheThrProTyrAsnIleValIleLeuLeuAsnThrPheGlnG 270
QY 993 ATTCTTTGGCTGAATAATTCAGTAGCTCTACAGGTGGACCAAGCTATCGAGTGAC 1052
Db 270 uPhePheCysLeuSerAsnGluSerThrSerGlnLeuAspGlnAlaThrGlnValTh 290
QY 1053 AGAGACTCTTGGGATGACGCTGCTCATCAACCCCATCATCTATGCTTTGTCGGGA 1112
Db 290 rGluThrLeuGlyMetThrHisCysCysIleAsnProIleIleIleIleIleIle 310
QY 1113 GAAGTTCAGAACTACTCTTAGTCTCTTCCAAAAGACATTCGCCAAGGCTTCGAA 1172
Db 310 uLysPheArgTyrLeuSerValPhePheArgLysHisIleThrLysArgPheCysLy 330
QY 1173 ATGCTGTCTATTTCCAGAGAGGCTCCGAGCGAGCAAGCTCAGTTTACACCGATC 1232
Db 330 sGlnCysProValPheTyrArgGluThrValAspGlyValThrSerThrAsnThrPro 350
QY 1233 CACTGGGAGCAGGAAATATCTGTGGGCTTG 1263
Db 350 rThrGlyGluGlnGluValSerAlaGlyLeu 360

RESULT 15

PCT-US95-00476-4
; Sequence 4, Application PC/TUS9500476
; GENERAL INFORMATION:
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: MAMMALIAN MONOCYTE CHEMOATTRACTANT
; TITLE OF INVENTION: PROTEIN RECEPTORS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Robbins, Berliner & Carson
; STREET: 201 N. Figueroa Street, 5th Floor
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90012-2628
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/00476
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Berliner, Robert
; REGISTRATION NUMBER: 20,121
; REFERENCE/DOCKET NUMBER: 5555-291
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 310-977-1001
; TELEFAX: 310-977-1003
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 360 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US95-00476-4

Alignment Scores:
Pred. No.: 9,96e-137 Length: 360
Score: 1331.00 Matches: 258
Percent Similarity: 83.76% Conservative: 36
Best Local Similarity: 73.50% Mismatches: 46
Query Match: 50.65% Indels: 11
DB: 5 Gaps: 3

[illegible]

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 3, 2003, 18:49:55 ; Search time 64.7773 Seconds
(without alignments)
4506.655 Million cell updates/sec

Title: US-09-938-703-3
Perfect score: 2628
Sequence: 1 GATTTCCTCCACAGACCA.....AGTAGATTAGATCGAATTC 1442

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 383519 seqs, 101223694 residues
Total number of hits satisfying chosen parameters: 767038

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEU-framer_n2p.model -DEV-xlp
-Q/cgn2_1/USPRO.spool/US0938703/runat_03062003_161410_23015/app_query.fasta_1.1870
-DB-PublishedApplications_AA -QFMT=fastan -SUFFIX=rapb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS-human40.cdi -LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US0938703 -CGN_1_1_24 -runat_03062003_161410_23015
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications_AA.*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pcp.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pcp.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pcp.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pcp.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pcp.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pcp.*
7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pcp.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pcp.*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pcp.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pcp.*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pcp.*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pcp.*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pcp.*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----|-------------------------------------|
| 1 | 1745.5 | 66.4 | 352 | 9 | US-10-086-814-1 Sequence 1, Appli |
| 2 | 1745.5 | 66.4 | 352 | 9 | US-09-734-221A-14 Sequence 14, Appl |
| 3 | 1745.5 | 66.4 | 352 | 10 | US-09-759-841-2 Sequence 2, Appli |
| 4 | 1745.5 | 66.4 | 352 | 10 | US-09-813-653-15 Sequence 15, Appl |

| | | | | | | |
|----|--------|------|-----|----|-------------------|--------------------|
| 5 | 1745.5 | 66.4 | 352 | 10 | US-09-796-202-1 | Sequence 1, Appli |
| 6 | 1745.5 | 66.4 | 352 | 10 | US-09-938-719-5 | Sequence 5, Appli |
| 7 | 1745.5 | 66.4 | 352 | 10 | US-09-939-226-5 | Sequence 5, Appli |
| 8 | 1745.5 | 66.4 | 352 | 10 | US-09-938-703-5 | Sequence 5, Appli |
| 9 | 1745.5 | 66.4 | 352 | 12 | US-10-106-623-2 | Sequence 2, Appli |
| 10 | 1739.5 | 66.2 | 352 | 10 | US-09-813-653-17 | Sequence 17, Appli |
| 11 | 1737.5 | 66.1 | 352 | 9 | US-10-232-686-2 | Sequence 2, Appli |
| 12 | 1737.5 | 66.1 | 352 | 10 | US-09-725-285-2 | Sequence 22, Appli |
| 13 | 1737.5 | 66.1 | 352 | 10 | US-09-779-879A-22 | Sequence 22, Appli |
| 14 | 1737.5 | 66.1 | 352 | 10 | US-09-779-880A-22 | Sequence 22, Appli |
| 15 | 1737.5 | 66.1 | 352 | 10 | US-09-195-662A-2 | Sequence 2, Appli |
| 16 | 1737.5 | 66.1 | 352 | 10 | US-09-339-912A-2 | Sequence 2, Appli |
| 17 | 1737.5 | 66.1 | 352 | 10 | US-09-502-783A-2 | Sequence 2, Appli |
| 18 | 1730.5 | 65.8 | 352 | 10 | US-09-779-879A-2 | Sequence 2, Appli |
| 19 | 1730.5 | 65.8 | 352 | 10 | US-09-779-880A-2 | Sequence 2, Appli |
| 20 | 1711.5 | 65.1 | 352 | 12 | US-10-106-623-20 | Sequence 20, Appli |
| 21 | 1332 | 50.7 | 360 | 10 | US-09-131-827A-20 | Sequence 20, Appli |
| 22 | 1331 | 50.6 | 347 | 10 | US-09-104-792-3 | Sequence 3, Appli |
| 23 | 1331 | 50.6 | 360 | 10 | US-09-131-827A-2 | Sequence 2, Appli |
| 24 | 1317 | 50.1 | 360 | 10 | US-09-938-719-7 | Sequence 7, Appli |
| 25 | 1317 | 50.1 | 360 | 10 | US-09-939-226-7 | Sequence 7, Appli |
| 26 | 1317 | 50.1 | 360 | 10 | US-09-938-703-7 | Sequence 7, Appli |
| 27 | 1190.5 | 45.3 | 344 | 9 | US-10-232-686-9 | Sequence 9, Appli |
| 28 | 1190.5 | 45.3 | 344 | 10 | US-09-779-879A-9 | Sequence 9, Appli |
| 29 | 1190.5 | 45.3 | 344 | 10 | US-09-779-880A-9 | Sequence 9, Appli |
| 30 | 1122 | 42.7 | 215 | 10 | US-09-938-719-6 | Sequence 6, Appli |
| 31 | 1122 | 42.7 | 215 | 10 | US-09-939-226-6 | Sequence 6, Appli |
| 32 | 1122 | 42.7 | 215 | 10 | US-09-938-703-6 | Sequence 6, Appli |
| 33 | 1105 | 42.0 | 329 | 10 | US-09-725-285-9 | Sequence 9, Appli |
| 34 | 1105 | 42.0 | 329 | 10 | US-09-195-662A-9 | Sequence 9, Appli |
| 35 | 1105 | 42.0 | 329 | 10 | US-09-339-912A-9 | Sequence 9, Appli |
| 36 | 1105 | 42.0 | 329 | 10 | US-09-502-783A-9 | Sequence 9, Appli |
| 37 | 1008 | 38.4 | 375 | 9 | US-10-219-834-78 | Sequence 78, Appli |
| 38 | 1003.5 | 38.2 | 355 | 10 | US-09-961-068-1 | Sequence 1, Appli |
| 39 | 1003.5 | 38.2 | 355 | 10 | US-09-960-547-1 | Sequence 1, Appli |
| 40 | 976.5 | 37.2 | 355 | 10 | US-09-938-719-9 | Sequence 9, Appli |
| 41 | 976.5 | 37.2 | 355 | 10 | US-09-939-226-9 | Sequence 9, Appli |
| 42 | 976.5 | 37.2 | 355 | 10 | US-09-938-703-9 | Sequence 9, Appli |
| 43 | 958 | 36.5 | 184 | 10 | US-09-938-719-4 | Sequence 4, Appli |
| 44 | 958 | 36.5 | 184 | 10 | US-09-939-226-4 | Sequence 4, Appli |
| 45 | 958 | 36.5 | 184 | 10 | US-09-938-703-4 | Sequence 4, Appli |

ALIGNMENTS

RESULT 1
US-10-086-814-1
; Sequence 1, Application US/10086814
; Publication No. US20030092632A1
; GENERAL INFORMATION:
; APPLICANT: Dragic, Tatjana
; APPLICANT: Olson, William C.
; TITLE OF INVENTION: SULFATED CCR5 PEPTIDES FOR HIV-1 INFECTION
; FILE REFERENCE: 61010-AB-1
; CURRENT APPLICATION NUMBER: US/10/086.814
; CURRENT FILING DATE: 2002-02-28
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-086-814-1

Alignment Scores:
Pred. No.: 1.73e-158
Score: 1745.50
Percent Similarity: 96.60%
Best Local Similarity: 96.32%
Query Match: 66.42%
DB: 9
Matches: 352
Conservative: 1
Mismatches: 0
Indels: 12
Gaps: 1

US-09-938-703-3 (1-1442) x US-10-086-814-1 (1-352)

| | | | |
|----|------|---|------|
| QY | 240 | ATGGATTATCAAGTGTCAAGTCCCAATCTATGACATCAATATTATTATACATCGGAGCCCTGC | 299 |
| Db | | | |
| Db | 1 | MetAspTyrGlnValSerSerProIleTyrAspIleAsnTyrTyrThrSerGluProCys | 20 |
| QY | 300 | CAAAAACAAATGTGAAGCAAAATCGACGCCGCTCCCTGGCTCGGCTCTACTCACTGCTGGT | 359 |
| Db | | | |
| Db | 21 | GlnIysIleAsnValLysGlnIleAlaAlaArgLeuLeuProIleTyrSerLeuVal | 40 |
| QY | 360 | TTCACTTTGGTTTTTGGGGCAACATCGTGTGATCACTCATCTCATCTGATAAATCAAAAGG | 419 |
| Db | | | |
| Db | 41 | PheIlePheGlyPheValGlyAsnMetLeuValIleLeuIleLeuIleAsnCysLysArg | 60 |
| QY | 420 | CTGAAGAGCATGATGACATCTACTCTCAACTCGCCATCTCTGACCTGTTTTCCTT | 479 |
| Db | | | |
| Db | 61 | LeuLysSerMetIrispIleTyrLeuLeuAsnLeuAlaIleSerAspLeuPhePheLeu | 80 |
| QY | 480 | CTTACTGTCCCTTCTGGGCTCACTATGTGTGGCGCCAGTGGGACTTTGGAAATACAATG | 539 |
| Db | | | |
| Db | 81 | LeuThrValProPheTrpAlaHisTyrAlaAlaLagIlnTrpAspPheGlyAsnThrMet | 100 |
| QY | 540 | TGTCACACTTGCACAGGCTCTATTTATAGCTTCTTCTGGAATCTTCTTCATCAATC | 599 |
| Db | | | |
| Db | 101 | CysGlnLeuLeuThrGlyLeuTyrPheIleGlyPhePheSerGlyIlePhePheIleIle | 120 |
| QY | 600 | CTCCTGCACAAATGATAGGTACTCGTGGCTGTGCTCCATGCTGTGTTGCTTTAAAGCCAGG | 659 |
| Db | | | |
| Db | 121 | LeuLeuThrIleAspArgTyrLeuAlaValHisAlaValPheAlaLeuLysAlaArg | 140 |
| QY | 660 | ACGGTCACCTTTGGGGTGTGCACAGTGTGATCACTTGGTGGTGGCTGTGTTGGCTGT | 719 |
| Db | | | |
| Db | 141 | ThrValThrPheGlyValValThrSerValIleThrTrpValValAlaValPheAlaSer | 160 |
| QY | 720 | CTCCAGGAATCATCTTACCAGATCTCAAAAGAGGCTCTTCATTACACCTGCAGCTCT | 779 |
| Db | | | |
| Db | 161 | LeuProGlyIleIlePheThrArgSerGlnLysGluGlyLeuHisTyrThrCysSerSer | 180 |
| QY | 780 | CATTTTCCTATAC-----ATTAAGATAGTACAT | 806 |
| Db | | | |
| Db | 181 | HisPheProTyrSerGlnTyrGlnPheTrpLysAsnPheGlnThrLeuLysIleValI | 200 |
| QY | 807 | CTTGGGGTGTGCTCCCGCTGCTGTCAATGTCATCTGCTACTCGGAATCCCTAAAC | 866 |
| Db | | | |
| Db | 200 | eLeuGlyLeuValLeuProLeuValMetValIleCysTyrSerGlyIleLeuLysTh | 220 |
| QY | 867 | TCGTCTCGGTCGAAATGACAGAGGACACAGGCGTCGAGGCTTATCTTCACCAT | 926 |
| Db | | | |
| Db | 220 | rLeuLeuArgCysArgAsnGluLysArgHisArgAlaValArgLeuIlePheThrI | 240 |
| QY | 927 | CATGATGTTTATTTCTCTCTGGGTCCTCTACACATTTGCTCTCTCGACACCTT | 986 |
| Db | | | |
| Db | 240 | eMetIleValTyrPheLeuPheTrpAlaProTyrAsnIleValLeuLeuLeuAsnThrPh | 260 |
| QY | 987 | CCAGGAATCTTTGGGCTCAATAATGCGATAGCTCTTAACAGGTTGGACCAGCTATGCA | 1046 |
| Db | | | |
| Db | 260 | eGlnGluPhePheGlyLeuAsnAsnCysSerSerAsnArgLeuAspGlnAlaMetG | 280 |
| QY | 1047 | GSTGACAGACATCTTGGGATCAGCAGCTGCTGCGATCAACCCCATCATCTATGCTTGT | 1106 |
| Db | | | |
| Db | 280 | nValThrGluThrLeuGlyMetThrHisCysCysIleAsnProIleIleTyrAlaPheVa | 300 |
| QY | 1107 | CGGGGAGAAGTTCAGAAATCACTCTTAGTCTTCTTCCAAAAGCATTTGCCAAGCCTT | 1166 |
| Db | | | |
| Db | 300 | IgLYGluLysPheArgAsnTyrLeuLeuValPhePheGlnLysHisIleAlaLysArgPh | 320 |
| QY | 1167 | CTGCAATAGCTGTGTTCTATTTTCCAGCAAGAGGCTCCCGAGCGACAGCACTCACTTTAC | 1226 |
| Db | | | |
| Db | 320 | eCysLysCysCysSerIlePheGlnGlnGluAlaProGluArgAlaSerSerValTyrTh | 340 |
| QY | 1227 | CCGATCCCATCGGGACAGGAATATCTGTGGCTTG | 1263 |
| Db | | | |
| Db | 340 | rArgSerThrGlyGluGlnGluIleSerValGlyLeu | 352 |

```

RESULT 2
US-09-734-221A-14
; Sequence 14, Application US/09734221A
; Publication No. US20030096221A1
; GENERAL INFORMATION:
APPLICANT: LITTMAN, DAN R.
; DENG, HONGKUI
; ELLMEIER, WILFRIED
; LANDAU, NATHANIEL R.
; LIU, RONG
;
; TITLE OF INVENTION: G-COUPLED RECEPTORS ASSOCIATED WITH
; MACROPHAGE-TROPIC HIV, AND DIAGNOSTIC AND THERAPEUTIC
; USES THEREOF
;
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
ADDRESSEE: David A. Jackson, Esq.
; FLOOR
; STREET: 411 Hackensack Ave, Continental Plaza, 4th
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/734,221A
; FILING DATE: 11-Dec-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/666,020
; FILING DATE: 19-JUN-1996
APPLICATION NUMBER: US 08/227,319
; FILING DATE: 13-APR-1994
; ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 1049-1-004 N2
; TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
LENGTH: 352 amino acids
; TYPE: amino acid
STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-09-734-221A-14

Alignment Scores:
Pred. No.: 1,73e-158 Length: 352
Score: 1745.50 Matches: 340
Percent Similarity: 96.60% Conservative: 1
Best Local Similarity: 96.32% Mismatches: 0
Query Match: 66.42% Indels: 12
DB: 9 Gaps: 1

US-09-938-703-3 (1-1442) x US-09-734-221A-14 (1-352)

Qy 240 ATGGATTATCAAGTGCATGCTCCAAATCTATGACATCAATATTATATACATCGAGCCCTGC 299
;
Db 1 MetAspTyr-GlnValSerProIleTyrAspIleAsnTyrTyrThrSerGluProCys 20
;
Qy 300 CAAAATCAATGTCGACCAATCGAGCCCTCTCCCTCCGCTCTACTCTACTCTACTGCTG 359
;
Db 21 GlnLysIleAsnValLysGlnIleAlaAlaArgLeuLeuProLeuTyrSerLeuVal 40
;

```


Db 220 rLeuLeuArgCysArgAsnGluLysLysArgHisArgAlaValArgLeuIlePheThrIl 240
QY 927 CATGATGTTGTTATTTCTCTCTGGGCTCCCTACACATGTTCTCTCTCCCTGAACACCTT 986
Db 240 eMetIleValTyPheLeuPheThrPalaProTyrrAsnIleValLeuLeuLeuAsnThrPh 260
QY 987 CCAGGAATCTTTGGCTGAATAATGCACTAGTCTTAAACAGGTTGGACCAAGCTATGCA 1046
Db 260 eGlnGluPhePheGlyLeuAsnAsnCysSerSerAsnArgLeuAspGlnAlaMetGl 280
QY 1047 GGTGACAGACACTCTGGGATGACGACACTGCTGCATCAACCCCATCATGAGCTTTGT 1106
Db 280 nValThrGluThrLeuGlyMetThrHisCysCysIleAsnProIleIleTyAlaPheVa 300
QY 1107 CGGGGAGAAGTTCAGAACTACCTCTTAGTCTCTTCCAAAAGCACATGCCCACAGCTT 1166
Db 300 lGlyGluLysPheArgAsnTyLeuLeuValPhePheGlnLysHisIleAlaLysArgPh 320
QY 1167 CTGCAATGCTGTTCTATTCTTCCAGCAAGAGGCTCCGAGCGAGCAAGCTCAGTTTACAC 1226
Db 320 eCysLysCysCysSerIlePheGlnGlnGluAlaProGluArgAlaSerSerValTyTh 340
QY 1227 CGGATCCACTGGGACGACGAGAAATATCTGTGGCTTG 1263
Db 340 rArgSerThrGlyGluGlnGluIleSerValGlyLeu 352

RESULT 4

US-09-813-653-15
; Sequence 15, Application US/09813653
; Patent No. US20020064770A1
; GENERAL INFORMATION:
; APPLICANT: Nestor, John
; APPLICANT: Wilson, Carol
; APPLICANT: See, Raymond
; APPLICANT: Tan Hehir, Christina
; TITLE OF INVENTION: Binding Compounds and Methods For Identifying Binding Compounds
; FILE REFERENCE: CMS-005
; CURRENT APPLICATION NUMBER: US/09/813,653
; CURRENT FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: US 60/190,946
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/190,996
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/191,299
; PRIOR FILING DATE: 2000-03-21
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 15
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-813-653-15

Alignment Scores:

| | | | |
|------------------------|-----------|---------------|-----|
| Pred. No.: | 1-73e-158 | Length: | 352 |
| Score: | 1745.50 | Matches: | 340 |
| Percent Similarity: | 96.60% | Conservative: | 1 |
| Best Local Similarity: | 96.32% | Mismatches: | 0 |
| Query Match: | 66.42% | Indels: | 12 |
| DB: | 10 | Gaps: | 1 |

US-09-938-703-3 (1-1442) x US-09-813-653-15 (1-352)

QY 240 ATGGATTATCAAGTGTCAAGTCCAAATCTATGACATCAATATTATATACATCGGAGCCCTGC 299
Db 1 MetaAspTyrrGlnValSerSerProIleTyrrAspIleAsnTyrrThrSerGluProCys 20
QY 300 CAARAATCAATGTGAAGCAATCGAGCCGCCCTCTGCTCGCTCTACTACTGCTGTG 359
Db 21 GlnIleAsnValLysGlnIleAlaAlaArgLeuLeuProIleTyrrSerLeuVal 40
QY 360 TTCATCTTTGTTTGTGGGCAACATCTGCTGGTCACTCCCTCACTCACTGATAAATCGCAAGG 419

Db 41 PheIlePheGlyPheValGlyAsnMetLeuValIleLeuLeuLeuAsnCysLysArg 60
QY 420 CTGAGACATGACTGACATCTACTGCTCACTGACCTGACCTGACCTGACCTGACCTGACCT 479
Db 61 LeuLysSerMetThrAspIleTyrrLeuLeuAsnLeuAlaIleSerAspLeuPhePheLeu 80
QY 480 CTTACTGTCCCTTCTCTGGGCTCATTGCTGCGCCGACGCTGGGACTTTGGAATCAATG 539
Db 81 LeuThrValProPheThrPalaHisTyrrAlaAlaGlnTrpAspPheGlyAsnThrMet 100
QY 540 TGTCAACTCTTGAACAGGCTCTATTTTATAGGCTTCTCTCTGGAATCTTCTTCATCATC 599
Db 101 CysGlnLeuLeuThrGlyLeuTyrrPheIleGlyPhePheSerGlyIlePhePheIleIle 120
QY 600 CTCTCGCAATCGATAGTACTGCTGCTGCTCCATGCTGCTGCTGCTGCTGCTGCTGCTG 659
Db 121 LeuLeuThrIleAspArgTyrrLeuAlaValHisAlaValPheAlaLeuLysAlaArg 140
QY 660 ACGGTCACTTTGGGCTGTGACAAAGTGTGATCATTGGGTGGTGGTGGTGGTGGTGGT 719
Db 141 ThrValThrPheGlyValValThrSerValIleThrTriPValValAlaValPheAlaSer 160
QY 720 CTCCAGGAATCATCTTTTACCAGATCTCAAAAAGAGGCTTCATTCACCTGCACTCT 779
Db 161 LeuProGlyIleIlePheThrArgSerGlnLysGluGlyLeuHisTyrrThrCysSerSer 180
QY 780 CATTTCCTCCATAC-----NTTAAAGATGACTCAT 806
Db 181 HisPheProTyrrSerGlnTyrrGlnPheThrIlyAsnPhGlnThrLeu-LysIleValIl 200
QY 807 CTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 866
Db 200 eLeuGlyLeuValLeuProLeuLeuValMetValIleCysTyrrSerGlyIleLeuLysTh 220
QY 867 TCTGCTCGGTGTCGAAATGAGAAGAGGACAGGCTGTGAGGCTTATCTTCACCAT 926
Db 220 rLeuLeuArgCysArgAsnGluLysLysArgHisArgAlaValArgLeuIlePheThrIl 240
QY 927 CATGATGTTTATTTCTCTCTGGGCTCCCTACACATGTTCTCTCTCTGACACCTT 986
Db 240 eMetIleValTyrrPheLeuPheThrPalaProTyrrAsnIleValLeuLeuLeuAsnThrPh 260
QY 987 CCAGGAATCTTTGGCTGAATAATGCACTAGTCTTAAACAGGTTGGACCAAGCTATGCA 1046
Db 260 eGlnGluPhePheGlyLeuAsnAsnCysSerSerAsnArgLeuAspGlnAlaMetGl 280
QY 1047 GGTGACAGACACTCTTGGGATGACGACACTGCTGCATCAACCCCATCATGCTTTGT 1106
Db 280 nValThrGluThrLeuGlyMetThrHisCysCysIleAsnProIleIleTyrrAlaPheVa 300
QY 1107 CGGGGAGAAGTTCAGAACTACCTCTTAGTCTTCTTCCAAAAGCACATTCGCAAGCTT 1166
Db 300 lGlyGluLysPheArgAsnTyrrLeuLeuValPhePheGlnLysHisIleAlaLysArgPh 320
QY 1167 CTGCAATGCTGTTCTATTCTTCCAGCAAGAGGCTCCGAGCGAGCAAGCTCAGTTTACAC 1226
Db 320 eCysLysCysCysSerIlePheGlnGlnGluAlaProGluArgAlaSerSerValTyTh 340
QY 1227 CGGATCCACTGGGACGACGAGAAATATCTGTGGCTTG 1263
Db 340 rArgSerThrGlyGluGlnGluIleSerValGlyLeu 352

RESULT 5

US-09-796-202-1
; Sequence 1, Application US/09796202
; Patent No. US20020068813A1
; GENERAL INFORMATION:
; APPLICANT: Dragic, Tatjana
; APPLICANT: Olson, William
; TITLE OF INVENTION: SULFATED CCR5 PEPTIDES FOR HIV-1 INFECTION
; FILE REFERENCE: 2048/61010/JPW/SHS
; CURRENT APPLICATION NUMBER: US/09/796,202

; CURRENT FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 352
; TYPE: PRT
; ORGANISM: human
US-09-796-202-1

Alignment Scores:

Pred. No.: 1-73e-158 Length: 352
Score: 1745.50 Matches: 340
Percent Similarity: 96.60% Conservatives: 1
Best Local Similarity: 96.32% Mismatches: 0
Query Match: 96.42% Indels: 12
DB: 10 Gaps: 1

US-09-938-703-3 (1-1442) x US-09-796-202-1 (1-352)

```
QY 240 ATGGATTATCAAGTGTCAAGTCCAAATCTATGACATCAATATTATATACATCGGAGCCCTGC 299
Db 1 MetAspTyrGlnValSerSerProIleTyrAspIleAsnTyrTyrThrSerGluProCys 20
QY 300 CAAAAAATCAATGTGAAGCAAAATCGAGCCGCTCTGCTCGCTCGCTCTACTCTACTCTGGTG 359
Db 21 GlnLysIleAsnValLysGlnIleAlaAlaArgLeuLeuProLeuTyrSerLeuVal 40
QY 360 TTCATCTTTGGTTTGGGCAACATGCTGTGTCATCTCATCTCATCTGATAACTGCAAAAGS 419
Db 41 PheIlePheGlyPheValGlyAsnMetLeuValIleLeuIleLeuIleAsnCysLysArg 60
QY 420 CTGAAGCATGACTGACATCTACCTGCTCAACCTGCGCAATCTCTGACTGTTCCTT 479
Db 61 LeuLysSerMetThrAspIleTyrLeuLeuAsnLeuAlaIleSerAspLeuPheLeu 80
QY 480 CTTACTGTCCTCTGCTGCTACTATGCTGCGCCAGTGGGACTTGGAAATACAAATG 539
Db 81 LeuThrValProPheIlePheAlaHisTyrAlaAlaGlnIlePhePheIleAsnMet 100
QY 540 TGTCAACTCTGACAGGCTCTATTTATPAGGCTTCTCTGGAATCTTCTTCATCATC 599
Db 101 CysGlnLeuLeuThrGlyLeuTyrPheIleGlyPhePheSerGlyIlePhePheIle 120
QY 600 CTCCTGACATGATAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 659
Db 121 LeuLeuThrIleAspTyrLeuAlaValAlaValHisAlaValPheAlaLeuLysAlaArg 140
QY 660 ACGTCACTTTGGGCTGTTGACAACTGTGATCATCTGGTGGTGGTGGTGGTGGTGGTCT 719
Db 141 ThrValThrPheGlyValValThrSerValIleThrIleThrIleThrIleThrIle 160
QY 720 CTCCTGAGTAATCATCTTACAGATCTCAAAAGAGGTCTTCATPACACCTGAGCTCT 779
Db 161 LeuProGlyIleIlePheThrArgSerGlnLysGluGlyLeuHisTyrThrCysSer 180
QY 780 CATTTTCCATAC-----ATTAAAGATAGTCAT 806
Db 181 HisPheProTyrSerGlnTyrGlnPheTyrLysAsnPheGlnThrLeuLysIleValI 200
QY 807 CTTGGGCTGGTCTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 866
Db 200 eLeuGlyLeuValLeuProLeuValMetValIleCysTyrSerGlyIleLeuLysTh 220
QY 867 TCTGCTTCGGTGTGGAATGAGAGAGAGAGGACAGGCTGTGAGGCTTACTTTCACCAT 926
Db 220 rLeuLeuArgCysArgAsnGluLysLysArgHisArgAlaValArgLeuIlePheThr 240
QY 927 CATGATGTTTATTTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 986
Db 240 eMetIleValTyrPheLeuPheThrAlaProTyrAsnIleValLeuLeuLeuAsnThrPh 260
QY 987 CCAGGAATCTTTGGCTGAATAATTGCAAGTACTTCAACAGGTGGGAGTGGCAAGCTATGCA 1046
```

```
Db 260 eGlnGluPhePheGlyLeuAsnAsnCysSerSerSerAsnArgLeuAspGlnAlaMetGI 280
QY 1047 GGTGACAGACACTCTTTGGGATGAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1106
Db 280 nValThrGluThrLeuGlyMetThrHisCysCysIleAsnProIleTyrAlaPheVa 300
QY 1107 CGGGGAGAAGTTCAGAACTACCTCTTAGTCTCTTCCAAAGACAAATGCGCAACGCTT 1166
Db 300 lGlyGluLysPheArgAsnTyrLeuLeuValPhePheGlnLysHisIleAlaLysArgPh 320
QY 1167 CTGCAATGCTGTTCTATTTCAGCAAGAGGCTCCCGAGCGAGCAGCTCAGTTCAC 1226
Db 320 eCysLysCysSerIlePheGlnGlnGluAlaProGluArgAlaSerSerValTyrTh 340
QY 1227 CCGATCCACTGGGAGCAGGAAATATCTGTGGGCTTG 1263
Db 340 rArgSerThrGlyGluGlnGluIleSerValGlyLeu 352
```

RESULT 6

US-09-938-719-5
; Sequence 5, Application US/09938719
; Patent No. US20020106742A1
; GENERAL INFORMATION:

APPLICANT: SAMSON, MICHEL
PARMENTIER, MARC
VASSART, GILBERT

LIBERT, FREDERICK
TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR
AND NUCLEIC ACID MOLECULES ENCODING SAID RECEPTOR

NUMBER OF SEQUENCES: 17

CORRESPONDENCE ADDRESS:

ADDRESSEE: Knobbe, Martens, Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: CA

COUNTRY: U.S.A.

ZIP: 92660

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA: US/09/938, 719
FILING DATE: 24-Aug-2001
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/626,939

FILING DATE: 27-JULY-2000

ATTORNEY/AGENT INFORMATION:

NAME: Altman, Daniel E

REGISTRATION NUMBER: 34,115

REFERENCE/DOCKET NUMBER: <Unknown>

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 352 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 5:

US-09-938-719-5

Alignment Scores:

Pred. No.: 1-73e-158 Length: 352
Score: 1745.50 Matches: 340
Percent Similarity: 96.60% Conservatives: 1
Best Local Similarity: 96.32% Mismatches: 0
Query Match: 96.42% Indels: 12
DB: 10 Gaps: 1

US-09-938-703-3 (1-1442) x US-09-938-719-5 (1-352)

QY 240 ATGGATTATCAAGTGTCAAGTCCAAATCTATGACATCAATATTATATACATCGGAGCCCTGC 299

Db 1 MetAspTyrGlnValSerSerProIleTyrAspIleAsnTyrThrSerGluProCys 20
QY 300 CAAATAATCAATGTGAGCAATCGAGCGCCCTCTGCTCGCTCTACTCTACTGCTG 359
Db 21 GlnLysIleAsnValLysGlnIleAlaAlaArgLeuLeuProLeuTyrSerLeuVal 40
QY 360 TTCATCTTTGTTTGTGGCAACATGCTGCTCCTCATCTCTGATACCTGCAAAAGG 419
Db 41 PheIlePheGlyPheValGlyAsnMetLeuValIleLeuLeuIleAsnCysLysArg 60
QY 420 CTGAAGAGCATGATGATCATCTACTCTCAACCTGGCCATCTCTGACCTGTTTTCCT 479
Db 61 LeuLysSerMetThrAspIleTyrLeuLeuAsnLeuAlaIleSerAspLeuPheLeu 80
QY 480 CTACTGCTCCCTTCTGGGCTCACTATGCTGCGCCGAGTGGAGCTTGGAAATACATG 539
Db 81 LeuThrValProPheTrpAlaHisTyrAlaAlaAlaGlnTrpAspPheGlyAsnThrMet 100
QY 540 TGTCAACTCTTGACAGGCTCTATTTATATAGGCTTCTCTGGAATCTTCTCATCATC 599
Db 101 CysGlnLeuLeuThrGlyLeuTyrPheIleGlyPhePheSerGlyIlePhePheIle 120
QY 600 CTCCTGACATCATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 659
Db 121 LeuLeuThrIleAspArgTyrLeuAlaValAlaValHisAlaValPheAlaLeuLysAlaArg 140
QY 660 ACAGTTCACCTTTGGGCTGTCGACAGTGTGATCACTGCTGGTGGTGGTGGTGGTCT 719
Db 141 ThrValThrPheGlyValValThrSerValIleThrTrpValValAlaValPheAlaSer 160
QY 720 CTCGAGAAATCATCTTACAGATCTCAAGATCTCAAAAGAGCTCTCATCTACCTGCACTCT 779
Db 161 LeuProGlyIleIlePheThrArgSerGlnLysGluGlyLeuHisTyrThrCysSerSer 180
QY 780 CATTTTCCATAC-----ATTAAAGATAGTCAT 806
Db 181 HisPheProIleTyrSerGlnTyrGlnPheTrpLysAsnPheGlnThrLeuLysIleValI 200
QY 807 CTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 866
Db 200 eLeuGlyLeuValLeuProLeuLeuValMetValIleCysTyrSerGlyLeuLeuLysTh 220
QY 867 TCTGCTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 926
Db 220 rLeuLeuArgCysArgAsnGlnLysLysArgHisArgAlaValArgLeuIlePheThrI 240
QY 927 CATGATGTTTATTTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 986
Db 240 eMetIleValTyrPheLeuPheTrpAlaProTyrAsnIleValLeuLeuLeuAsnThrPh 260
QY 987 CCAGAAATCTTTGGCTGAATATGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1046
Db 260 eGlnGluPhePheGlyLeuAsnAsnCysSerSerSerAsnArgLeuAspGlnAlaMetG 280
QY 1047 GGTGACAGACTCTGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1106
Db 280 nValThrGluThrLeuGlyMetThrHisCysCysIleAsnProIleIleTyrAlaPheVa 300
QY 1107 CGGGGAGAGTTCAGAACTACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1166
Db 300 lGlyGluLysPheArgAsnTyrLeuLeuValPhePheGlnLysHisIleAlaLysArgPh 320
QY 1167 CTGCAATGCTGCTTATTTTCCAGAGAGGCTCCGAGGAGGAGGAGGAGGAGGAGGAG 1226
Db 320 eCysLysCysCysSerIlePheGlnGlnGluAlaProGluArgAlaSerSerValTyrTh 340
QY 1227 CCGATCCACTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1263
Db 340 rArgSerThrGlyGluGlnGluIleSerValGlyLeu 352

RESULT 7

US-09-939-226-5

; Sequence 5, Application US/09939226
; Patent No. US20020110805A1
; GENERAL INFORMATION:
; APPLICANT: SAMSON, MICHEL
; PARMENTIER, MARC
; VASSART, GILBERT
; LIBERT, FREDERICK
; TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR
; AND NUCLEIC ACID MOLECULES ENCODING SAID RECEPTOR
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/939/226
; FILING DATE: 24-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/626,939
; FILING DATE: 2000-07-27
; ATTORNEY/AGENT INFORMATION:
; NAME: Altman, Daniel E
; REGISTRATION NUMBER: 34,115
; REFERENCE/DOCKET NUMBER: <Unknown>
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 352 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-939-226-5

Alignment Scores:

| | | | |
|------------------------|-----------|---------------|-----|
| Pred. No.: | 1,73e-158 | Length: | 352 |
| Score: | 1745.50 | Matches: | 340 |
| Percent Similarity: | 96.60% | Conservative: | 1 |
| Best Local Similarity: | 96.32% | Mismatches: | 0 |
| Query Match: | 66.42% | Indels: | 12 |
| DB: | 10 | Gaps: | 1 |

US-09-938-703-3 (1-1442) x US-09-939-226-5 (1-352)

| | | | |
|----|-----|--|-----|
| QY | 240 | ATGATATTCAAGTGTCAAGTCCCAATCTATGACATCAATATTATATGACGAGCCCTGC | 299 |
| Db | 1 | MetAspTyrGlnValSerSerProIleTyrAspIleAsnTyrThrSerGluProCys | 20 |
| QY | 300 | CAAAATAATCAATGTGAGCAATCGAGCGCCCTCTGCTCGCTCTACTCTACTGCTG | 359 |
| Db | 21 | GlnLysIleAsnValLysGlnIleAlaAlaArgLeuLeuProLeuTyrSerLeuVal | 40 |
| QY | 360 | TTCATCTTTGTTTGTGGCAACATGCTGCTCCTCATCTCTGATACCTGCAAAAGG | 419 |
| Db | 41 | PheIlePheGlyPheValGlyAsnMetLeuValIleLeuLeuIleAsnCysLysArg | 60 |
| QY | 420 | CTGAAGAGCATGATGATCATCTACTCTCAACCTGGCCATCTCTGACCTGTTTTCCT | 479 |
| Db | 61 | LeuLysSerMetThrAspIleTyrLeuLeuAsnLeuAlaIleSerAspLeuPheLeu | 80 |
| QY | 480 | CTACTGCTCCCTTCTGGGCTCACTATGCTGCGCCGAGTGGAGCTTGGAAATACATG | 539 |
| Db | 81 | LeuThrValProPheTrpAlaHisTyrAlaAlaAlaGlnTrpAspPheGlyAsnThrMet | 100 |
| QY | 540 | TGTCAACTCTTGACAGGCTCTATTTATATAGGCTTCTCTGGAATCTTCTCATCATC | 599 |

| | | | |
|----|------|--|------|
| | | LeuValLeuProLeuLeuValMetValIleCysTyrSerGlyIleLeuLysThr | 220 |
| Db | 200 | eLeuGlyLeuValLeuProLeuLeuValMetValIleCysTyrSerGlyIleLeuLysThr | 220 |
| | | TCTGTTCGGTGCGAAATGAGAGAAGACGACAGGGCTGTGAGCCTTACTTCAACAT | 926 |
| QY | 867 | TCTGTTCGGTGCGAAATGAGAGAAGACGACAGGGCTGTGAGCCTTACTTCAACAT | 926 |
| | | rLeuLeuArgCysArgAsnGluLysLysArgHisArgAlaValArgLeullePheThrIle | 240 |
| Db | 220 | rLeuLeuArgCysArgAsnGluLysLysArgHisArgAlaValArgLeullePheThrIle | 240 |
| | | CATGATTGTTATTTTTCTCTTCCTGGGCCTCCCTACAACATGTCTTCCTCGAACACCTT | 986 |
| QY | 927 | CATGATTGTTATTTTTCTCTTCCTGGGCCTCCCTACAACATGTCTTCCTCGAACACCTT | 986 |
| | | eMetIleValTyrPheLeuPheMetPAlaProTyrAsnIleValLeuLeuLeuAsnThrPh | 260 |
| Db | 240 | eMetIleValTyrPheLeuPheMetPAlaProTyrAsnIleValLeuLeuLeuAsnThrPh | 260 |
| | | CCAGGAATCTTTGGCCTGAATAATTGCAGTAGCTCTTAACAGTTGGACCACAGCTATGCA | 1046 |
| QY | 987 | CCAGGAATCTTTGGCCTGAATAATTGCAGTAGCTCTTAACAGTTGGACCACAGCTATGCA | 1046 |
| | | eGlnGluPhePheGlyLeuAsnAsnCysSerSerSerAsnArgLeuaspGlnAlaMetGl | 280 |
| Db | 260 | eGlnGluPhePheGlyLeuAsnAsnCysSerSerSerAsnArgLeuaspGlnAlaMetGl | 280 |
| | | GTTGACAGACACTCTGGGAGGACGACACTGTGTCATCAACCCCATCATCTATGCCTTGT | 1106 |
| QY | 1047 | GTTGACAGACACTCTGGGAGGACGACACTGTGTCATCAACCCCATCATCTATGCCTTGT | 1106 |
| | | nValThrGluThrLeuGlyMetThrHisCysGlyLeasNPrIleIleTyrAlaPheVa | 300 |
| Db | 280 | nValThrGluThrLeuGlyMetThrHisCysGlyLeasNPrIleIleTyrAlaPheVa | 300 |
| | | CGGGGAGAAGTCTCAGAAACTACCTCTTAGCTTCTTCCAAGACACATGCGAAGCGTT | 1166 |
| QY | 1107 | CGGGGAGAAGTCTCAGAAACTACCTCTTAGCTTCTTCCAAGACACATGCGAAGCGTT | 1166 |
| | | lGlyGluLysPheArgAsnTyrLeuLeuValPhePheGlnLysHisIleAlalysArgph | 320 |
| Db | 300 | lGlyGluLysPheArgAsnTyrLeuLeuValPhePheGlnLysHisIleAlalysArgph | 320 |
| | | CTGCAAAATGCTGTCTATTATTTCCACGAAGAGGCTCCGAGCGGACGAAGCTCAGTTTACAC | 1226 |
| QY | 1167 | CTGCAAAATGCTGTCTATTATTTCCACGAAGAGGCTCCGAGCGGACGAAGCTCAGTTTACAC | 1226 |
| | | eCysLysCysCysSerIlePheGlnGlnAlaproGluArgAlaSerSerValTyrTh | 340 |
| Db | 320 | eCysLysCysCysSerIlePheGlnGlnAlaproGluArgAlaSerSerValTyrTh | 340 |
| | | CGANTCACTGGGGAGCAGGAATATCTGTGGGCTTG | 1263 |
| QY | 1227 | CGANTCACTGGGGAGCAGGAATATCTGTGGGCTTG | 1263 |
| | | rArgSerThrGlyGluGlnGluIleSerValGlyLeu | 352 |
| Db | 340 | rArgSerThrGlyGluGlnGluIleSerValGlyLeu | 352 |

RESULT 9

```

US-10-106-623-2
: Sequence 2, Application US/10106623
: Patent No. US20020150888A1
: GENERAL INFORMATION:
: APPLICANT: Gray, Patrick W.
:             Schweickart, Vicki L.
:             Raport, Carol J.
: TITLE OF INVENTION: Chemokine Receptor Materials and Methods
: NUMBER OF SEQUENCES: 20
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
: STREET: 6300 Sears Tower, 233 S. Wacker Drive
: CITY: Chicago
: STATE: Illinois
: COUNTRY: USA
: ZIP: 60606
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/10106.623
: FILING DATE: 26-Mar-2002
: CLASSIFICATION: <Unknown>
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/771,276
: FILING DATE: <Unknown>
: ATTORNEY/AGENT INFORMATION:
: NAME: No. US20020150888A1and, Greta E.
: REGISTRATION NUMBER: 35,302
: REFERENCE/DOCKET NUMBER: 27866/33670
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 312-474-6300
: TELEFAX: 312-474-0448
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 352 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein

```


| | | |
|------|--|------|
| 101 | CysGlnLeuLeuThrGlyLeuTyPheIleGlyPheSerGlyIlePhePheIleIle | 120 |
| 600 | CTCCGACAACTCATAGTACTGGCTGCSCCATCATCTGTTTGGTTTAAAGCCAGG | 659 |
| 121 | LeuLeuThrIleAspArgTyLeuAlaValAlaValHisAlaValPheAlaLeuLysAlaArg | 140 |
| 660 | ACGGTCACCTTGGGTGGTGACAAAGTGTATCACTTGGGTGGTGGCTGTGTTCGCT | 719 |
| 141 | ThrValThrPheGlyValValThrSerValIleThrTrpValValAlaValPheAlaSer | 160 |
| 720 | CTCCGAGGAATCATCTTTACAGATCTCAAAAGAGAGGTCTTCATACACCTGCAGCTCT | 779 |
| 161 | LeuProGlyIleIlePheThrArgSerGlnLysGluGlyLeuHisTyThrCysSerSer | 180 |
| 780 | CATTTTCCATAC-----ATTAAGATAGTACAT | 806 |
| 181 | HisPheProTySerGlnTyGlnPheTrpLysAsnPheGlnThrLeu-LysIleValI | 200 |
| 807 | CTTGGGGCTGGTCTCCGCGCTGTGCATGTCATCTACTCGGAATCTCAAAAC | 866 |
| 200 | eLeuGlyLeuValLeuProLeuValMetValIleCysTySerGlyIleLeuLysTh | 220 |
| 867 | TCTGTCTCGGTGCGAAATGAGAAGAGGACAGGCGTGAGGCTTATTCATTCACAT | 926 |
| 220 | rLeuLeuArgCysArgAsnGlnLysArgHisArgAlaValArgLeuIlePheThrI | 240 |
| 927 | CATGATCTTTTATTTCTTCTCTGGGTCCCTACAAATTCCTCTCTCTGACACCTT | 986 |
| 240 | eMetIleValTyPheLeuPheTrpAlaProTyArgAsnIleValLeuLeuAsnThrPh | 260 |
| 987 | CCAGGAATCTTGGCGCTCAATATTCGATAGTCATCAACAGGTGGACCAAGCTATGCA | 1046 |
| 260 | eGlnGluPhePheGlyLeuAsnAsnCysSerSerAsnArgLeuAspGlnAlaMetG | 280 |
| 1047 | GSTGACAGACATCTTGGGATGAGCATCTGTCATCAACCCATCATCTATGCGCTTGT | 1106 |
| 280 | nValThrGlnThrLeuGlyMetThrHisCysCysIleAsnProIleIleTyAlaPheVa | 300 |
| 1107 | CGGGAGAGGTTCAGAACTACCTTTAGTCTTCTTCCAAAGCACATGCGCAAGCGTT | 1166 |
| 300 | lGlyGlnLysPheArgAsnTyLeuLeuValPhePheGlnLysHisIleAlaLysArgPh | 320 |
| 1167 | CTCAAAATCGTGTTCTATTTTCCAGCAAGAGCTCCGAGCGGACCAAGCTCAGTTTACAC | 1226 |
| 320 | eCysLysCysCysSerIlePheGlnGlnAlaProGluArgAlaSerSerValTyTh | 340 |
| 1227 | CCGATCCACTGGGAGCAGGAAATATCTGTGGCGTTG | 1263 |
| 340 | rArgSerThrGlyGlnGlnGlnIleSerValGlyLeu | 352 |

| | | | |
|----|-----|---|-----|
| QY | 240 | ATGGATTATCAAGTGTCAAGTCCCATCTATGACATCAATTATTATACATCGGAGCCCTGC | 299 |
| Db | 1 | | |
| Db | 1 | MeCaspIyGlnValSerProIleTyAspIleAsnTyThrSerGluProCys | 20 |
| QY | 300 | CAAAAAATCAATGTGAAGCAATCGCAGCCGCGCTCCCTGCCTCGCTTACTACTCGGTG | 359 |
| Db | 21 | | |
| Db | 21 | GlnLysIleAsnValLysGlnIleAlaAlaArgLeuLeuProLeuTySerLeuVal | 40 |
| QY | 360 | TTCATCTTTGGTTTTTGGGGCAACATCGCTGGTCACTCCATCCCTGATAAAGTCAAAAGG | 419 |
| Db | 41 | | |
| Db | 41 | PheIlePheGlyPheValGlyAsnMetLeuValIleLeuIleGlnIleAsnCysIleArg | 60 |
| QY | 420 | CTGACAGCATGACTGACATCTACCTGCTCAACCTGGCGCATCTCTGACCTGTTTTCCCTT | 479 |
| Db | 61 | | |
| Db | 61 | LeuLysSerMetThrAspIleTyLeuLeuAsnLeuAlaIleSerAspLeuPhePheLeu | 80 |
| QY | 480 | CTTACTGTCCCTCTCTGGGCTCAGTATGCTGGCGGCCAGTGGGACTTTGGAAATACAAATG | 539 |
| Db | 81 | | |
| Db | 81 | LeuThrValProPheThrAlaHisTyAlaAlaAlaGlnTrpaspPheGlyAsnThrMet | 100 |
| QY | 540 | TGTCAACTCTTGACAGCGGCTCATTTTATAGGCTTCTCTCGAATCTCTTCATCATC | 599 |

```

RESULT 11
US-10-232-686-2
; Sequence 2, Application US/10232686
; Publication No. US20030023044A1
; GENERAL INFORMATION:
; APPLICANT: Li, Yi
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Human G-Protein Chemokine Receptor (OCRS) HDGKR10
; FILE REFERENCE: 1498.115000N
; CURRENT APPLICATION NUMBER: US/10/232,686
; CURRENT FILING DATE: 2002-09-03
; PRIOR APPLICATION NUMBER: 09/339,912
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 09/195,662
; PRIOR FILING DATE: 1998-11-18
; PRIOR APPLICATION NUMBER: 08/466,343
; PRIOR FILING DATE: 1995-06-06
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 2
; LENGTH: 352
; TYPE: PRT

```


ORGANISM: Homo sapiens
US-10-232-686-2

Alignment Scores:

Pred. No.: 1,01e-157 Length: 352
Score: 1737.50 Matches: 339
Percent Similarity: 96.32% Conservative: 1
Best Local Similarity: 96.03% Mismatches: 1
Query Match: 66.11% Indels: 12
DB: 9 Gaps: 1

US-09-938-703-3 (1-1442) x US-10-232-686-2 (1-352)

```
QY 240 ATGGATTATCAAGTGTCAAGTCCAACTATGACATCAATATTATATACATCGGAGCCCTGC 299
Db 1 MetAspTyrGlnValSerProIleTyrAspIleAsnTyrThrSerGluProCys 20
QY 300 CAAAATCAATGTGAGCAATCCAGCCGCGCTCTGCTCCGCTCTACTCATCTGCTG 359
Db 21 GlnLysIleAsnValLysGlnIleAlaAlaArgLeuLeuProLeuTyrSerLeuVal 40
QY 360 TTCACTTTGGTTTGTGGCAACATGCTGCTCATCTCATCTGATAAATGCAAAAGG 419
Db 41 PheIlePheGlyPheValGlyAsnMetLeuValIleLeuIleAsnCysLysArg 60
QY 420 CTGAGAGCATGACTGACATCTACTGCTCAACCTGGCCATCTCTGACCTGTTTTCCTT 479
Db 61 LeuLysSerMetThrAspIleTyrLeuLeuAsnLeuAlaIleSerAspLeuPheLeu 80
QY 480 CTTACTGTCCCTCTGGCTCCTACTATGCTGGCCGCGAGTGGAGCTTGGAAATCAATG 539
Db 81 LeuThrValProPheThrPheIleHisTyrAlaAlaAlaGlnTrpAspPheGlyAsnThrMet 100
QY 540 TGTCAACTCTGTGACGGCTCATTTATATAGGCTTCTCTGGAATCTCTCATCATC 599
Db 101 CysGlnLeuLeuThrGlyLeuTyrPheIleGlyPheSerGlyIlePhePheIleIle 120
QY 600 CTCCTGACATCGATAGTACCTGGCTGTGCTCCATGCTGTGTTGCTTTAAAGCCAGG 659
Db 121 LeuLeuThrIleAspArgTyrLeuAlaValAlaHisAlaValPheAlaLeuLysAlaArg 140
QY 660 AGGTCACCTTTGGGTGGTGACAGTGTGATCCTGCTGGTGGTGGTGTGCTGCT 719
Db 141 ThrValThrPheGlyValValThrSerValIleThrTrpValValAlaValPheAlaSer 160
QY 720 CTCACAGATCATCTTTACAGATCTCAAAAGAGGTCTTCATACCTGCGAGCTCT 779
Db 161 LeuProGlyIleIlePheThrArgSerGlnLysGluGlyLeuHisTyrThrCysSerSer 180
QY 780 CATTTCCTCATAC-----ATTAAGATAGTCAT 806
Db 181 HisPheProTyrSerGlnTyrGlnPheThrPheLysAsnPheGlnThrLeu-LysIleValI 200
QY 807 CTTGGGGTGGTCTCCGCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 866
Db 200 eLeuGlyLeuValLeuProLeuLeuValMetValIleCysTyrSerGlyIleLeuLysTh 220
QY 867 TCTGCTCGGTGTCGAATGAGAGAGACAGCGGCTGTGAGGCTTATCTTCACCAT 926
Db 220 rLeuLeuArgCysArgAsnGluLysLysArgHisArgAlaValArgLeuIlePheThrIL 240
QY 927 CATGATTCTTTATTTCTCTCTGGCTCCCTACACATGCTGCTCTCTCTCTCTCTCTCT 986
Db 240 eMetIleValTyrPheLeuPheThrPalAProTyrAsnIleValLeuLeuLeuAsnThrPh 260
QY 987 CCAGGAATCTTTGGCTGAATAATGCTAGTCTTAACAGGTGGACCAAGCTATGCA 1046
Db 260 eGlnGluPhePheGlyLeuAsnAsnCysSerSerAsnArgLeuAspGlnAlaMetG 280
QY 1047 GTGACAGACTCTTGGAGTACGCGACTCTGCTGATCAACCCCATCATCTATGCTTTGT 1106
Db 280 nValThrGluThrLeuGlyMetThrHisCysCysIleAsnProIleIleTyrAlaPheVa 300
```

```
QY 1107 CGGGAGAGTTCAGAACTACTCTTAGTCTTCTTCCAAAGACATTCGCAAGCTT 1166
Db 300 lGlyGluLysPheArgAsnTyrLeuLeuValPhePheGlnLysHisIleAlaLysArgPh 320
QY 1167 CTCGAAATGCTGTTCTATTTCAGCAAGAGCTCCGAGCGAGCAAGCTCAGTTTACAC 1226
Db 320 eCysLysCysCysSerIlePheGlnGlnGluAlaProGluArgAlaSerSerValTyrTh 340
QY 1227 CCGATCCACTGGGAGCAGGAGAAATATCTGTGGGTG 1263
Db 340 rArgSerThrGluGluGlnGlnIleSerValGlyLeu 352
```

RESULT 12

US-09-725-285-2
Sequence 2, Application US/09725285
Patent No. US20010000241A1
GENERAL INFORMATION:
APPLICANT: Li, Yi
TITLE OF INVENTION: Antibodies to Human G-Protein Chemokine Receptor HDGMR10
FILE REFERENCE: 1488, 1150003
CURRENT APPLICATION NUMBER: US/09/725,285
CURRENT FILING DATE: 2000-11-29
PRIOR APPLICATION NUMBER: 09/339,912
PRIOR FILING DATE: 1999-06-25
PRIOR APPLICATION NUMBER: 09/195,662
PRIOR FILING DATE: 1998-11-18
PRIOR APPLICATION NUMBER: 08/466,343
PRIOR FILING DATE: 1995-06-06
NUMBER OF SEQ ID NOS: 9
SOFTWARE: PatentIn version 3.0
SEQ ID NO 2
TYPE: PRT
LENGTH: 352
ORGANISM: Artificial Sequence: Genomic
FEATURE:
OTHER INFORMATION: Deduced Amino Acid Sequence
US-09-725-285-2

Alignment Scores:

Pred. No.: 1,01e-157 Length: 352
Score: 1737.50 Matches: 339
Percent Similarity: 96.32% Conservative: 1
Best Local Similarity: 96.03% Mismatches: 1
Query Match: 66.11% Indels: 12
DB: 10 Gaps: 1

US-09-938-703-3 (1-1442) x US-09-725-285-2 (1-352)

```
QY 240 ATGGATTATCAAGTGTCAAGTCCAACTATGACATCAATATTATATACATCGGAGCCCTGC 299
Db 1 MetAspTyrGlnValSerProIleTyrAspIleAsnTyrThrSerGluProCys 20
QY 300 CAAAATCAATGTGAGCAATCCAGCCGCGCTCTGCTCCGCTCTACTCATCTGCTG 359
Db 21 GlnLysIleAsnValLysGlnIleAlaAlaArgLeuLeuProLeuTyrSerLeuVal 40
QY 360 TTCACTTTGGTTTGTGGCAACATGCTGCTCATCTCATCTGATAAATGCAAAAGG 419
Db 41 PheIlePheGlyPheValGlyAsnMetLeuValIleLeuIleAsnCysLysArg 60
QY 420 CTGAGAGCATGACTGACATCTACTGCTCAACCTGGCCATCTCTGACCTGTTTTCCTT 479
Db 61 LeuLysSerMetThrAspIleTyrLeuLeuAsnLeuAlaIleSerAspLeuPheLeu 80
QY 480 CTTACTGTCCCTCTGGCTCCTACTATGCTCCGCGCGAGTGGAGCTTGGAAATCAATG 539
Db 81 LeuThrValProPheThrPheIleHisTyrAlaAlaAlaGlnTrpAspPheGlyAsnThrMet 100
QY 540 TGTCAACTCTTGGAGTACGCGACTCTGCTGATCAACCCCATCATCTATGCTTTGT 1106
Db 101 CysGlnLeuLeuThrGlyLeuTyrPheIleGlyPheSerGlyIlePhePheIleIle 120
```


Score: 1737.50 Matches: 339
Percent Similarity: 96.32% Conservative: 1
Best Local Similarity: 96.03% Mismatches: 1
Query Match: 66.11% Indels: 12
DB: 10 Gaps: 1

US-09-938-703-3 (1-1442) x US-09-195-662A-2 (1-352)

```
QY 240 ATGGATTATCAAGTGTCAAGTCCAAATCTATGACATCAATATTATATACATCGAGCCCTGC 299
Db 1 MetAspTyrGlnValSerSerProIleTyrAspIleAsnTyrTyrThrSerGluProCys 20
QY 300 CAAAAATCAATGACGACCAATCGGAGCCCGCTCTCCCTCCCTCTACTCAGTGGTG 359
Db 21 GlnLysIleAsnValysGlnIleAlaAlaArgLeuLeuProIleuTyrSerLeuVal 40
QY 360 TTCATCTTTGGTTTGGGCAACATGCTGTGGTCTCATCTCTCATCTGATTAACCTGCAAAAGG 419
Db 41 PheIlePheGlyPheValGlyAsnMetLeuValIleLeuIleLeuIleAsnCysLysArg 60
QY 420 CTGAGAGCATGACGACATCTACTGCTCAACCTGGCCATCTGACCTGTTTTCCTT 479
Db 61 LeuLysSerMetThrAspIleTyrLeuLeuAsnLeuAlaIleSerAspLeuPhePheLeu 80
QY 480 CTTACTGTCCTCTCTGGGCTCACTATGTCGCCCGGAGGACTTTGGAAATACAAATG 539
Db 81 LeuThrValProPheTrpAlaHisTyrAlaAlaAlaGlnTrpAspPheGlyAsnThrMet 100
QY 540 TGTCAACTCTTGACAGGCGCTATTATTTATAGGCTTCTCTGGAATCTTCTCATCATC 599
Db 101 CysGlnLeuLeuThrGlyLeuTyrPheIleGlyPheSerGlyIlePhePheIleIle 120
QY 600 CTCCTGACAAATGATGAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 659
Db 121 LeuLeuThrIleAspArgTyrLeuAlaValValHisAlaValPheAlaLeuLysAlaArg 140
QY 660 ACGCTCACCTTTGGGCTGCTGACAGTGTGATCATCTGGTGGTGGTGGTGGTGGTGGTCT 719
Db 141 ThrValThrPheGlyValValThrSerValIleThrTrpValValAlaValPheAlaSer 160
QY 720 CTCGCCAGGAATCATCTTTACCAGATCTCAAAAGAAGTCTTCAATACACCTGACGCTCT 779
Db 161 LeuProGlyIleIlePheThrArgSerGlnLysGluGlyLeuHisTyrThrCysSerSer 180
QY 780 CATTTTCCATAC-----ATTAAAGATAGTCAT 806
Db 181 HisPheProTyrSerGlnTyrGlnPheTrpLysAsnPheGlnThrLeu-LysIleValI 200
QY 807 CTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 866
Db 200 eLeuGlyLeuValLeuProLeuLeuValMetValIleCysTyrSerGlyIleLeuLysTh 220
QY 867 TCTGCTTCGGTCTCAAAATGAGAGAGAGGACAGGCTGTGAGGCTTATCTTTCACCAT 926
Db 220 rLeuLeuArgCysArgAsnGluLysLysArgHisArgAlaValArgLeuIlePheThrI 240
QY 927 CATGATTTTATTTTCTCTCTGGGCTCCCTACACATTCCTCTCTCTCTCTCTCTCTCT 986
Db 240 eMetIleValTyrPheLeuPheTrpAlaProTyrAsnIleValLeuLeuLeuAsnThrPh 260
QY 987 CCAGGAATCTTTGGCTGCAATATTGCAAGTAGCTCTAAGAGCTTGGACCAAGCTATGCA 1046
Db 260 eGlnGluPhePheGlyLeuAsnAsnCysSerSerSerAsnArgLeuAspGlnAlaMetGl 280
QY 1047 GGTGACAGAGATCTTGGGATGAGCAGCTGCTGATCATCAACCCCATCATCTATGCTTTGT 1106
Db 280 nValThrGluThrLeuGlyMetThrHisCysCysIleAsnProIleIleTyrAlaPheVa 300
QY 1107 CGGGGAGAGAGTTCAGAACTACCTCTTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1166
Db 300 lGlyGluLysPheArgAsnTyrLeuLeuValPhePheGlnLysHisIleAlaLysArgPh 320
QY 1167 CTGCAAAATGCTGTTCTATTTCATTTTCCAGCAAGAGGCTCCGAGGAGCAAGCTCAGTTTACAC 1226
```

```
Db 320 eCysLysCysCysSerIlePheGlnGlnGlnAlaProGluArgAlaSerSerValTyrTh 340
QY 1227 CGATCCACTGGGGAGCAGCAAAATATCTGTGGGCTTG 1263
Db 340 rArgSerThrGluGluGlnGlnIleSerValGlyLeu 352
```

Search completed: June 3, 2003, 19:13:30
Job time : 83.7773 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 3, 2003, 18:42:09 ; Search time 50.6953 seconds
(without alignments)
5468.980 Million cell updates/sec

Title: US-09-938-703-3

Perfect score: 2828

Sequence: 1 GAATTCCTCCCAACAGACCA.....AGTAGATTAGATCCGAATTC 1442

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 566448

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_n2p_model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09938703/runat_03062003_161407_22855/app.query.fasta_1.1870
-DB=PIR73 -QFMT=Eastan -SUFFIX=rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosom62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09938703 -CGN_1_1_89 -runat_03062003_161407_22855 -NCPU=3
-NO_MAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPOP=0.5 -FGAPOP=6
-FGAPOP=7 -YGAPOP=10 -YGAPOP=0.5 -DELOP=6 -DELEXT=7

Database: PIR73.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|--------|--------------------|
| 1 | 1745.5 | 66.4 | 352 | A43113 | chemokine (C-C) re |
| 2 | 1331 | 50.6 | 360 | JC2443 | chemokine (C-C) re |
| 3 | 1191 | 45.3 | 374 | J38450 | chemokine (C-C) re |
| 4 | 1003.5 | 38.2 | 355 | A45177 | chemokine (C-C) re |
| 5 | 960 | 36.5 | 359 | J49341 | MIP-1 alpha recept |
| 6 | 953 | 36.3 | 355 | J49339 | macrophage inflam |
| 7 | 909 | 34.6 | 355 | G02436 | chemokine (C-C) re |
| 8 | 891.5 | 33.9 | 360 | A57160 | chemokine (C-C) re |
| 9 | 878.5 | 33.4 | 360 | JC4587 | chemokine (C-C) re |
| 10 | 797.5 | 30.3 | 383 | J55594 | G protein-coupled |
| 11 | 776 | 29.5 | 356 | J49340 | MIP-1 alpha recept |
| 12 | 716 | 27.2 | 355 | JC5067 | G protein-coupled |
| 13 | 688.5 | 26.2 | 354 | J58186 | probable G protein |
| 14 | 644 | 24.5 | 355 | JC4304 | orphan G protein-c |

| | | | | | | |
|----|-------|------|-----|---|--------|--------------------|
| 15 | 638.5 | 24.3 | 344 | 2 | JC5942 | chemokine receptor |
| 16 | 573.5 | 21.8 | 378 | 2 | B55735 | lymphocyte-specifi |
| 17 | 559.5 | 21.3 | 378 | 2 | A45680 | G protein-coupled |
| 18 | 556.5 | 21.2 | 378 | 2 | A55735 | G protein-coupled |
| 19 | 552.5 | 21.0 | 369 | 2 | JC5068 | G protein-coupled |
| 20 | 528.5 | 20.1 | 333 | 2 | J65989 | G protein-coupled |
| 21 | 511.5 | 19.5 | 360 | 2 | A53611 | interleukin-8 rece |
| 22 | 507 | 19.3 | 359 | 2 | A48921 | interleukin-8 rece |
| 23 | 505 | 19.2 | 355 | 2 | JQ1231 | interleukin-8 rece |
| 24 | 501.5 | 19.1 | 367 | 2 | JF0349 | interferon-inducib |
| 25 | 501 | 18.1 | 358 | 2 | A53752 | interleukin-8 rece |
| 26 | 498 | 18.9 | 350 | 2 | JN0621 | G protein-coupled |
| 27 | 492.5 | 18.7 | 350 | 2 | A39445 | interleukin-8 rece |
| 28 | 491.5 | 18.7 | 352 | 2 | G00048 | fusin (LESTRA) - c |
| 29 | 491.5 | 18.7 | 352 | 2 | A45747 | neuropeptide Y pep |
| 30 | 486 | 18.5 | 356 | 2 | S42096 | interleukin-8 rece |
| 31 | 483.5 | 18.4 | 353 | 2 | S28787 | neuropeptide Y pep |
| 32 | 468 | 17.8 | 374 | 2 | S42628 | G protein-coupled |
| 33 | 463 | 17.6 | 359 | 2 | A42656 | angiotensin II rec |
| 34 | 462 | 17.6 | 359 | 2 | JC2134 | angiotensin II rec |
| 35 | 460 | 17.5 | 359 | 2 | S44425 | angiotensin II rec |
| 36 | 459 | 17.5 | 359 | 2 | S15403 | angiotensin II rec |
| 37 | 458 | 17.4 | 359 | 2 | J51372 | angiotensin II rec |
| 38 | 457 | 17.4 | 359 | 2 | JH0621 | angiotensin II rec |
| 39 | 455.5 | 17.3 | 327 | 2 | S56162 | MDCR15 protein - h |
| 40 | 455.5 | 17.3 | 372 | 2 | S26667 | G protein-coupled |
| 41 | 454 | 17.3 | 359 | 2 | JQ1516 | angiotensin II rec |
| 42 | 453.5 | 17.3 | 374 | 2 | S32785 | G protein-coupled |
| 43 | 453 | 17.2 | 359 | 2 | JC1104 | angiotensin II rec |
| 44 | 452 | 17.2 | 359 | 2 | JC1194 | angiotensin II rec |
| 45 | 452 | 17.2 | 359 | 2 | A48857 | angiotensin II rec |

ALIGNMENTS

RESULT 1

A43113

chemokine (C-C) receptor 5 - human

N:Alternate names: C-C CKR-5; CCR5

C:Species: Homo sapiens (man)

C>Date: 12-Jul-1996 #sequence, revision 12-Jul-1996 #text change 20-Jun-2000

C:Accession: A43113; S71808; A58834; A58832; G02653; A58833

R:Samson, M.; Labbe, O.; Mollereau, C.; Vassart, M.

Biochemistry 35, 3362-3367, 1996

A>Title: Molecular cloning and functional expression of a new human CC-chemokine rec

A:Reference number: A43113; MUID:96241590; PMID:8639485

A:Accession: A43113

A:Molecule type: mRNA

A:Residues: 1-352 <SAM1>

A:Cross-references: GB:X91492; NID:q1262810; PIDN:CAA62796.1; PID:q1262811

R:Samson, M.; Libert, F.; Doranz, B.J.; Rucker, J.; Liesnard, C.; Farber, C.M.; Sarrac

M.; Imai, T.; Rana, S.; Yi, Y.; Smyth, R.J.; Collman, R.G.; Doms, R.W.; Vassart, G.;

Nature 382, 722-725, 1996

A>Title: Resistance to HIV-1 infection in caucasian individuals bearing mutant allele

A:Reference number: S71808; MUID:96345670; PMID:8751444

A:Accession: S71808

A>Status: nucleic acid sequence not shown; not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 182-206; 207-230 <SAM2>

A:Accession: A58834

A>Status: nucleic acid sequence not shown; not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-184; /IKDSHIGAGPAACHGHLGNPKNSASVK' <SAM3>

A:Cross-references: GB:X99393; NID:q1524062; PIDN:CAA67767.1; PID:q1524063

A>Note: this frameshift mutation results in a non-functional receptor but confers a c

nd may have had a selective advantage by conferring resistance to Yersinia plague inf

R:Combadiere, C.; Ahuja, S.K.; Tiffany, H.L.; Murphy, P.M.

J. Leukoc. Biol. 60, 147-152, 1996

A>Title: Cloning and functional expression of CC CKR5, a human monocyte CC chemokine

A:Reference number: A58832; MUID:96295970; PMID:8699119

A:Accession: A58832

A:Molecule type: mRNA

A:Residues: 1-352 <COM1>

A:Cross-references: GB:U57840; NID:91502408; PIDN:AA017071.1; PID:91502409
 A:Experimental source: clone 8, endotoxin-stimulated peripheral blood monocytes
 R:Combiadere, C.
 submitted to the EMBL Data Library, May 1996
 A:Reference number: H01541
 A:Accession: G02653
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-89, 'L', 91-352 <COM2>
 A:Cross-references: EMBL:U57840
 R:Raport, C.J.; Gosling, J.; Schweickart, V.L.; Gray, P.W.; Charo, I.F.
 J. Biol. Chem. 271, 17161-17166, 1996
 A:Title: Molecular cloning and functional characterization of a novel human CC chemokine
 A:Reference number: A58833; MUID:96291862; PMID:8663314
 A:Accession: A58833
 A:Molecule type: mRNA
 A:Residues: 1-352 <RAP>
 A:Cross-references: GB:U54994; NID:91457945; PIDN:AA050598.1; PID:91457946
 C:Comment: This is a receptor for chemokines MIP-1alpha (see PIR:A30574), MIP-1beta (see
 C:Comment: Macrophage- and dual-tropic strains of HIV-1 bind to a complex of chemokine
 C:Genetics:
 A:Gene: GDB:CMK3R5; CCR5; CCR-5; CC-CR-5; CKR5; ChemR13
 A:Cross-references: GDB:1230510; OMIM:601373
 A:Map position: 3p21-3p21
 C:Function:
 A:Description: G protein-coupled receptor for chemokines MIP-1alpha, MIP-1beta and RANTES
 A:Note: Probably acts to control granulocyte proliferation and differentiation
 C:Superfamily: vertebrate rhodopsin
 C:Keywords: AIDS; G protein-coupled receptor; glycoprotein; phosphoprotein; transmembran
 F:32-56/Domain: transmembrane #status predicted <TM1>
 F:67-87/Domain: transmembrane #status predicted <TM2>
 F:103-124/Domain: transmembrane #status predicted <TM3>
 F:143-166/Domain: transmembrane #status predicted <TM4>
 F:193-218/Domain: transmembrane #status predicted <TM5>
 F:236-257/Domain: transmembrane #status predicted <TM6>
 F:285-300/Domain: transmembrane #status predicted <TM7>
 F:20-269,101-178/disulfide bonds: #status predicted
 F:268/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:336,337,342/Binding site: phosphate (Ser) (covalent) #status predicted
 F:340,343/Binding site: phosphate (Thr) (covalent) #status predicted

Alignment Scores:

| Pred. No.: | 6,41e-146 | Length: | 352 |
|------------------------|-----------|---------------|-----|
| Score: | 1745.50 | Matches: | 340 |
| Percent Similarity: | 96.60% | Conservative: | 1 |
| Best Local Similarity: | 96.32% | Mismatches: | 0 |
| Query Match: | 66.42% | Indels: | 12 |
| DB: | 2 | Gaps: | 1 |

US-09-938-703-3 (1-1442) x A43113 (1-352)

QY 240 ATGGATTATCAAGTGTCAAGTCCAACTATGACATCAATTTATATCATCGAGGCGCTGC 299
 Db 1 MetAspYrGlnValSerProIleTyAspIleAsnTyThrSerGluProCys 20
 QY 300 CAAATAATCAATGTGAAGCAATCGCAGCGCCCTCGCTGCTCTACTCTACTCTGCTG 359
 Db 21 GlnYsIleAsnValYsGlnIleAlaAlaArgLeuLeuProProLeuYrSerLeuVal 40
 QY 360 TTCATCTTGGTTTGTGGGCAACATGCTGTCATCTCTCATCTCTGATAACTGCAAAAGG 419
 Db 41 PheIlePheGlyPheValGlyAsnMetLeuValIleLeuIleLeuLeuLeuAsnCysIleArg 60
 QY 420 CTGAGACATGACATGACATCTACCTGCTACCTGCGCATCTCTGACCTGTTTTCCTT 479
 Db 61 LeuYsSerMetThrAspIleTyLeuLeuAsnLeuAlaIleSerAspLeuPheLeu 80
 QY 480 CTTACTCTCCCTTCTGGGCTCACTATGCTGCGCCAGTGGAGCTTTGGAAATACATG 539
 Db 81 LeuThrValProPheTrpAlaHisTyAlaAlaAlaGlnIlePhePheGlyAsnThrMet 100
 QY 540 TGTCAACTCTTGACAGGCGCTATTATTATAGGCTTCTCTCTGGAATCTTCTTCATCATC 599
 |||||||

Db 101 CysGlnLeuLeuThrGlyLeuTyPheIleGlyPhePheSerGlyIlePhePheIleIle 120
 QY 600 CTCTCTGACATCGATAGTACCTGGCTGTGCTCATCTGCTGTTCGTTTAAAGCCAGG 659
 Db 121 LeuLeuThrIleAspA-gTyLeuAlaValValHisAlaValPheAlaLeuLysAlaArg 140
 QY 660 ACGTCACTCTGGGCTGGTGAAGTGTGATCACTTGGTGGTGGTGGTGGTGGTGGTGGT 719
 Db 141 ThrValThrPheGlyValValThrSerValIleThrTrpValAlaValPheAlaSer 160
 QY 720 CTCCAGGAATCATCTTTACAGATCTCAAAAGAGGCTTTCATTACACCTGACCTCT 779
 Db 161 LeuProGlyIleIlePheThrArgSerGlnYsGlnGlyLeuHisTyThrCysSerSer 180
 QY 780 CATTTCCTCAATC-----ATTAAGATAGTCAAT 806
 Db 181 HisPheProTyTrSerGlnTyGlnPheTrpLysAsnPheGlnThrLeuLysIleValI 200
 QY 807 CTTGGGCTGTGCTCTGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 866
 Db 200 eLeuGlyLeuValLeuProLeuLeuValMetValIleCysTyTrSerGlyIleLeuLysTh 220
 QY 867 TCTGCTTCTGCTGCTGAATGAGAGGACACAGGCTGTGAGGCTTATCTTCAACCT 926
 Db 220 rLeuLeuArgCysArgAsnGlnYsLysArgHisArgAlaValArgLeuIlePheThrI 240
 QY 927 CATGATTGTTTATTTCTCTGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 986
 Db 240 eMetIleValTyPheLeuPheTrpAlaProTyAsnIleValLeuLeuLeuAsnThrPh 260
 QY 987 CCAGGAATCTTCTGGCTGCTGAATATGCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCT 1046
 Db 260 eGlnGluPheGlyLeuAsnAsnCysSerSerSerSerSerSerSerSerSerSerSer 280
 QY 1047 GGTGACAGAGCTCTGGGATGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1106
 Db 280 nValThrGlnThrLeuGlyMetThrHisCysCysIleAsnProIleIleTyAlaPheVa 300
 QY 1107 CGGGGAGAGTTCAGAACTACTCTTGTAGTCTTCTTCCAAAGACATTCGCCAAGCTT 1166
 Db 300 lGlyGlnYsPheArgAsnTyLeuValPhePheGlnYsHisIleAlaLysArgPh 320
 QY 1167 CTCGAAATGCTGTTCTATTTTCACAGAGCTCCCGAGCGAGCAAGCTCAGTTTACAC 1226
 Db 320 eCysLysCysCysSerIlePheGlnGlnAlaProGluArgAlaSerSerValTyTh 340
 QY 1227 CCGATCCACTGGGAGCAGCAATATCTGTGGCTTG 1263
 Db 340 rArgSerThrGlyGlnGlnIleSerValGlyLeu 352

RESULT 2

JC2443
 chemokine (C-C) receptor 2, splice form B - human
 N:Alternate names: C-C CR-2; monocyte chemoattractant protein 1 receptor; monocyte c
 C:Species: Homo sapiens (man)
 C>Date: 21-Feb-1995 #sequence_revision 05-Apr-1995 #text_change 20-Jun-2000
 C:Accession: JC2443; 138463
 R:Yamagami, S.; Tokuda, Y.; Ishii, K.; Tanaka, H.; Endo, N.
 Biochem. Biophys. Res. Commun. 202, 1156-1162, 1994
 A:Title: cDNA cloning and functional expression of a human monocyte chemoattractant
 A:Reference number: A53477; MUID:94195821; PMID:8146186
 A:Accession: I38463
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-360 <RES>
 A:Molecule type: mRNA
 A:Residues: 1-360 <YAM>
 A:Cross-references: DBJ:D29984; NID:9531246; PIDN:BA06253.1; PID:9531247
 R:Charo, I.F.; Myers, S.J.; Herman, A.; Franci, C.; Connolly, A.J.; Coughlin, S.R.
 Proc. Natl. Acad. Sci. U.S.A. 91, 2752-2756, 1994
 A:Title: Molecular cloning and functional expression of two monocyte chemoattractant
 A:Reference number: A53477; MUID:94195821; PMID:8146186
 A:Accession: I38463
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-360 <RES>

A;Cross-references: EMBL:U03905; NID:9472557; PIDN:AAA19120.1; PID:9472558

C;Genetics:

A;Gene: GDB:CMKBR2

A;Cross-references: GDB:337364; OMTM:601267

A;Map position: 3p21-3p21

C;Superfamily: vertebrate rhodopsin

C;Keywords: alternative splicing; G protein-coupled receptor; glycoprotein; transmembran

F;43-70/Domain: transmembrane #status predicted <TM1>

F;81-100/Domain: transmembrane #status predicted <TM2>

F;115-136/Domain: transmembrane #status predicted <TM3>

F;154-178/Domain: transmembrane #status predicted <TM4>

F;207-226/Domain: transmembrane #status predicted <TM5>

F;244-268/Domain: transmembrane #status predicted <TM6>

F;297-309/Domain: transmembrane #status predicted <TM7>

F;14/Binding site: carbohydrate (Asn) (covalent) #status predicted

F;113-190/Disulfide bonds: #status predicted

Alignment Scores:

| | | | |
|------------------------|-----------|---------------|-----|
| Pred. No.: | 2,61e-109 | Length: | 360 |
| Score: | 1331.00 | Matches: | 258 |
| Percent Similarity: | 83.76% | Conservative: | 36 |
| Best Local Similarity: | 73.50% | Mismatches: | 46 |
| Query Match: | 50.65% | Indels: | 11 |
| DB: | 2 | Gaps: | 3 |

US-09-938-703-3 (1-1442) x JC2443 (1-360)

| | | | |
|----|-----|---|-----|
| QY | 234 | ARCAAGATGATATCAAGTCAAGTCCCAATCATCATCAATATTATATACATCGGAG | 293 |
| Db | 14 | AsnGluSerGlyGluGluValThrThr---PhePheAspTyrAspTyr-----GlyAla | 30 |
| QY | 294 | CCCTGCGCAAAATCATGTGACGAAATCGACCGCGCTCTCGCTCGCTCTACTCA | 353 |
| Db | 31 | ProCysHisLysPheAspValLysGlnIleGlyAlaGlnLeuProLeuTyrSer | 50 |
| QY | 354 | CTGGTGTTCATCTTGGTGTGGCAACATGCTGCTCATCTCATCTCATATAAATCG | 413 |
| Db | 51 | LeuValPheIlePheGlyPheValGlyAsnMetLeuValValleuLeuIleAsnCys | 70 |
| QY | 414 | AAAGGCTGAAGACGATGATGATGATGATGATGATGATGATGATGATGATGATG | 473 |
| Db | 71 | LysLysLeuLysCysLeuThrAspIleTyrLeuLeuAsnLeuAlaIleSerAspLeu | 90 |
| QY | 474 | TTCTTCTTACTGCTCCCTCTCTGGCTACTATGCTGCGCCGCTGGCTGGCTG | 533 |
| Db | 91 | PheLeuIleThrLeuProLeuThrPalaHisSerAlaAsnGluThrValPheGlyAsn | 110 |
| QY | 534 | ACAATGTGCACTCTTGACAGGCTCTATTATTAGGCTTCTCTCTGGAATCTTCT | 593 |
| Db | 111 | AlaMetCysLysLeuPheThrGlyLeuTyrHisIleGlyTyrPheGlyGlyLeuPhe | 130 |
| QY | 594 | ATCATCTCTGCAACATGATGATGATGATGATGATGATGATGATGATGATGATG | 653 |
| Db | 131 | IleIleLeuLeuThrIleAspArgTyrLeuAlaIleValHisAlaValPheAlaLeu | 150 |
| QY | 654 | GCCAGGACGCTCACTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT | 713 |
| Db | 151 | AlaArgThrValThrPheGlyValValThrSerValIleThrIlePheValAlaVal | 170 |
| QY | 714 | CGCTCTCTCCAGGAATCATCTTTACCAATCTCAAAAGAGGCTCTCATCATCTGC | 773 |
| Db | 171 | AlaSerValProGlyIleIlePheThrLysCysGlnLysGluAspSerValTyrVal | 190 |
| QY | 774 | AGC-----TCTCATTTCCATACATCAATCAATCAATCAATCAATCAATCAAT | 812 |
| Db | 191 | GlyProTyrPheProArgGlyTyrPalaAsnAsn-PheHisThrIleMetArgAsn | 210 |
| QY | 813 | GCTGCTCTCGCGCTCTCTCATGCTCATGCTCATGCTCATGCTCATGCTCATG | 872 |
| Db | 210 | YleuValLeuProLeuIleMetValIleCysTyrSerGlyIleLeuLysThrLeu | 230 |
| QY | 873 | TCGGTGTGAAATGAGAGAGAGGACAGGCTGTGAGGCTTATCTTCACCATCAT | 932 |

| | | | |
|----|------|--|------|
| Db | 230 | uArgCysArgAsnGluLysArgHisArgAlaValArgValIlePheThrIleMetI | 250 |
| QY | 933 | TGTTTATTTCTCTCTGCGCTCCCTACATGCTCTCTCTCTCTCTCTCTCTCTC | 992 |
| Db | 250 | eValTyrPheLeuPheThrProTyrAsnIleValIleLeuLeuAsnThrPheGln | 270 |
| QY | 993 | ATTTCTGGCTGAAATTCAGTACTGCTACAGGTTGACCAAGCTATGACAGTGAC | 1052 |
| Db | 270 | ubPhePheLeuSerAsnCysGluSerThrSerGlnLeuAspGlnAlaThrGlnVal | 290 |
| QY | 1053 | ACAGACTCTTGGATGACGACTGCTCATCAACCCCATCATCTATGCTCTGCGGGA | 1112 |
| Db | 290 | rGluThrLeuGlyMetThrHisCysCysIleAsnProIleIleTyrAlaPheVal | 310 |
| QY | 1113 | GAAATTCAGAACTACTCTTAGTCTTCTTCCAAAGCACATGCCAAACGCTCTG | 1172 |
| Db | 310 | uLysPheArgTyrLeuSerValPhePheArgLysHisIleThrLysArgPheCys | 330 |
| QY | 1173 | ATGCTGTCTTATTTCCAGCAAGAGCTCCGAGGAGCAAGCTCATGTTTACACCG | 1232 |
| Db | 330 | sGlnCysProValPheTyrArgGluThrValAlaPheGlyValThrSerThrAsn | 350 |
| QY | 1233 | CACCTGGGAGCAGGAATATCTGTGGCTTG | 1263 |
| Db | 350 | rThrGlyGluGlnGluValSerAlaGlyLeu | 360 |

RESULT 3

I38450

chemokine (C-C) receptor 2, splice form A - human

N;Alternate names: C-C CKR-2; monocyte chemoattractant protein 1 receptor; monocyte c

C;Species: Homo sapiens (man)

C;Date: 16-Feb-1996 #sequence_revision 16-Feb-1996 #text_change 13-Aug-1999

C;Accession: I38450

R;Charo, I.F.; Myers, S.J.; Herman, A.; Franci, C.; Connolly, A.J.; Coughlin, S.R.

A;Title: Molecular cloning and functional expression of two monocyte chemoattractant

A;Reference number: A53477; MUID:94195821; PMID:8146186

A;Accession: I38450

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-374 <RES>

A;Cross-references: EMBL:U03882; NID:9472555; PIDN:AAA19119.1; PID:9472556

C;Genetics:

A;Gene: GDB:CMKBR2

A;Cross-references: GDB:337364; OMTM:601267

A;Map position: 3p21-3p21

C;Superfamily: vertebrate rhodopsin

C;Keywords: alternative splicing; G protein-coupled receptor; glycoprotein; transmemb

F;44-68/Domain: transmembrane #status predicted <TM1>

F;79-99/Domain: transmembrane #status predicted <TM2>

F;115-136/Domain: transmembrane #status predicted <TM3>

F;154-178/Domain: transmembrane #status predicted <TM4>

F;208-226/Domain: transmembrane #status predicted <TM5>

F;244-265/Domain: transmembrane #status predicted <TM6>

F;292-309/Domain: transmembrane #status predicted <TM7>

F;14/Binding site: carbohydrate (Asn) (covalent) #status predicted

F;32-277,113-190/disulfide bonds: #status predicted

Alignment Scores:

| | | | |
|------------------------|----------|---------------|-----|
| Pred. No.: | 6.07e-97 | Length: | 374 |
| Score: | 1191.00 | Matches: | 234 |
| Percent Similarity: | 83.86% | Conservative: | 31 |
| Best Local Similarity: | 74.05% | Mismatches: | 34 |
| Query Match: | 45.32% | Indels: | 17 |
| DB: | 2 | Gaps: | 4 |

US-09-938-703-3 (1-1442) x I38450 (1-374)

QY 234 AACAGATGATATCAAGTCAAGTCCCAATCATCATCAATATTATATACATCGGAG

Db 14 AsnGluSerGlyGluGluValThrThr---PhePheAspTyrAspTyr-----GlyAla

QY 294 CCCTGCGCAAAATCATGTGACGAAATCGACCGCGCTCTCGCTCGCTCTACTCA


```

Db      31  ProCysHisLysPheAspValLysGlnIleGlyAlaGlnLeuLeuProLeuTyrSer 50
QY      354  CTGGTGTTCATCTTGGTTTGGGCAACATGCTGGTCATCCCTCATCCGATGAACAGCC 413
Db      51  LeuValPheIlePheGlyPheValGlyAsnMetLeuValValLeuLeuLeuLeuAsn 70
QY      414  AAAGSGCTGAAGAGCATGACTGACATCTACCTGCTCAACCTGGCCATCTCTGACCTGTT 473
Db      71  LysLysLeuLysCysLeuThrAspIleTyrLeuLeuAsnLeuAlaIleSerAspLeu 90
QY      474  TCTCTCTTACTGCTCCCTCTGGGCTCATCTATGCTGCGCCGACAGTGGAGCTTTGGAAT 533
Db      91  PheLeuIleThrLeuProLeuTrpAlaHisSerAlaAlaAsnGluTrpValPheGlyAsn 110
QY      534  ACAATGTGTCAACTGTGACAGGCTCTATTATAGCTCTCTCTCTGCAATCTCTTC 593
Db      111  AlaMetCysLysLeuPheThrGlyLeuTyrHisIleGlyTyrPheGlyGlyIlePhe 130
QY      594  ATCATCTCTCGACATGATAGTACCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 653
Db      131  IleIleLeuLeuThrIleAspArgTyrLeuAlaIleValHisAlaValPheAlaLeuLys 150
QY      654  GCACGAGGCTCACTTGGGCTGACAGTGTGATCATCTGCTGCTGCTGCTGCTGCTGCTGCT 713
Db      151  AlaArgThrValThrPheGlyValValThrSerValIleThrTrpLeuValAlaValPhe 170
QY      714  GCGTCTCCGACGATCATCTTACAGATCTCAAAAGAGGCTTTCATACACCTG 773
Db      171  AlaSerValProGlyIleIlePheThrLysCysGlnLysGluAspSerValTyrValCys 190
QY      774  AGC-----TCTCATTTTCCATACATTAAGATGATCATCTTGGG 812
Db      191  GlyProTyrPheProArgGlyTrpAsnAsn-PheHisThrIleMetArgAsnIleLeuG 210
QY      813  GCTGCTCTGCGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 872
Db      210  YleuValLeuProLeuLeuIleMetValIleCysTyrSerGlyIleLeuLysThrLeu 230
QY      873  TCGGTGTCGAATGAGAAGAGGACAGGCTGTGAGGCTTATCTTCCATCATCATGAT 932
Db      230  uArgCysArgAsnGluLysLysArgHisArgAlaValArgValIlePheThrIleMet 250
QY      933  TGTATTATTTCTCTTCTGGGCTCCCTACACATGCTGCTTCTCTGACACCTTCCAGGA 992
Db      250  eValTyrPheLeuPheThrTrpProTyrAsnIleValIleLeuLeuAsnThrPheGln 270
QY      993  ATTCTTGGCTGAATAATTGACGTAGCTCTAACAGGTGGACCAAGCTATGACAGTGCAC 1052
Db      270  uPhePheGlyLeuSerAsnCysGluSerThrSerGlnLeuAspGlnAlaThrGlnVal 290
QY      1053  AGAGACTCTGGGATGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1112
Db      290  rGlnThrLeuGlyMetThrHisCysCysIleAsnProIleTyrAlaPheValGlyG 310
QY      1113  GAAGTTCAGAACTACCTCTAGTCTTCTTCCAAAGACACATGCC 1158
Db      310  uLysPheArgSerLeuPhe-----HisIleAla 319

```

RESULT 4

A:Title: chemokine (C-C) receptor 1 - human
N:Alternate names: C-C CR-1; macrophage inflammatory protein-1-alpha receptor
C:Species: Homo sapiens (man)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 13-Aug-1999
C:Accession: A45177; 155671
R:Neote, K.; DiGregorio, D.; Mak, J. Y.; Horuk, R.; Schall, T. J.
Cell 72, 415-425, 1993
A:Title: Molecular cloning, functional expression, and signaling characteristics of a C-
cell 12, 415-425, 1993
A:Reference number: A45177; MUID:93161416; PMID:7679326
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA

A:Residues: 1-355 <NEO>
A:Cross-references: GB:L10918; NID:g292416; PIDN:AAA36543.1; PID:g292417
A:Experimental source: HL60 cells
A:Note: sequence extracted from NCBI backbone (NCBIP:124876)
R:Gao, J.
J. Exp. Med. 177, 1421-1427, 1993
A:Title: Structure and functional expression of the human macrophage inflammatory 1 a
A:Reference number: 155671; MUID:93240122; PMID:768036
A:Accession: 155671
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-355 <RES>
A:Cross-references: GB:L10918; NID:g292416; PIDN:AAA36543.1; PID:g292417
C:Genetics:
A:Gene: GDB:CMKRL; CMKR-1
A:Cross-references: GDB:138446; OMIM:601159
A:Map position: 3p21-3p21
C:Superfamily: vertebrate rhodopsin
C:Keywords: disulfide bond; G protein-coupled receptor; glycoprotein; phosphoprotein;
F:71-91/Domain: transmembrane #status predicted <TM1>
F:108-129/Domain: transmembrane #status predicted <TM2>
F:147-171/Domain: transmembrane #status predicted <TM3>
F:205-223/Domain: transmembrane #status predicted <TM4>
F:240-264/Domain: transmembrane #status predicted <TM5>
F:288-305/Domain: transmembrane #status predicted <TM6>
F:5/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:24-273,106-183/Disulfide bonds: #status predicted
F:345/Binding site: phosphate (Ser) (covalent) (by casein kinase II) #status predicted

Alignment Scores:

| | | | |
|------------------------|---------|---------------|-----|
| Pred. No.: | 2,2e-80 | Length: | 355 |
| Score: | 1003.50 | Matches: | 192 |
| Percent Similarity: | 71.31% | Conservative: | 59 |
| Best Local Similarity: | 54.55% | Mismatches: | 83 |
| Query Match: | 38.18% | Indels: | 18 |
| DB: | 2 | Gaps: | 5 |

US-09-938-703-3 (1-1442) x A45177 (1-355)

```

QY      243  GATTATCAAGTGTCAATCAATCTATGACATCAATATTATATACATGGAGCCCTCCCAA 302
Db      9  AspTyrAspThrThr-----GluPheAspTyrGlyAspAlaThrProCysGln 25
QY      303  AAAATCAATGATGAAGCAAAATCGACGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 362
Db      26  LysValAsnGluArgAlaPheGlyAlaGlnLeuLeuProLeuTyrSerLeuValPhe 45
QY      363  ATCTTTGGTTTGGGCAACATGCTGCTCATCTCTCATCTCTGATAAAGGCTG 422
Db      46  ValIleGlyLeuValGlyAsnIleLeuValValLeuValGlnTyrLysArgLeu 65
QY      423  AAGACGATGATGACATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 482
Db      66  LysAsnMetThrSerIleTyrLeuLeuAsnLeuAlaIleSerAspLeuLeuPhe 85
QY      483  ACTGTCCCTTCTGGGCTCACTATGCTGCCGCC---CAGTGGGACTTTGGAAATACATG 539
Db      86  ThrLeuProPheTrpIleAspTyrLysLeuLysAspTrpValPheGlyAspAlaMet 105
QY      540  TGTCAACTCTTGACAGGGCTCTATTTTATAGGCTTCTTCTGGAATCTTCTTCATCATC 599
Db      106  CysLysIleLeuSerGlyPheTyrTyrThrGlyLeuTyrSerGluIlePhePheIle 125
QY      600  CTCTCGAACAATGATAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 659
Db      126  LeuLeuThrIleAspArgTyrLeuAlaIleValHisAlaValPheAlaLeuArgAla 145
QY      660  ACGGTCACTTTGGGCTGGTGAAGTGTGATCATCTGCTGCTGCTGCTGCTGCTGCTGCT 719
Db      146  ThrValThrPheGlyValIleThrSerIleIleIleIleIleIleIleIleIleIle 165
QY      720  CTCGCCGAAGTATCTCTTTACCATCTCAAAAGAGGCTCTTATTACCTGCAGCTCT 779

```


| | | | |
|-----|----|--|-----|
| 279 | QY | TATTATATACGCGCGCTCGCCAAATAATCATGTGAAGCAATAATCGAGCCGCTCTCTG | 338 |
| 280 | QY | | |
| 281 | QY | | |
| 282 | QY | | |
| 283 | QY | | |
| 284 | QY | | |
| 285 | QY | | |
| 286 | QY | | |
| 287 | QY | | |
| 288 | QY | | |
| 289 | QY | | |
| 290 | QY | | |
| 291 | QY | | |
| 292 | QY | | |
| 293 | QY | | |
| 294 | QY | | |
| 295 | QY | | |
| 296 | QY | | |
| 297 | QY | | |
| 298 | QY | | |
| 299 | QY | | |
| 300 | QY | | |
| 301 | QY | | |
| 302 | QY | | |
| 303 | QY | | |
| 304 | QY | | |
| 305 | QY | | |
| 306 | QY | | |
| 307 | QY | | |
| 308 | QY | | |
| 309 | QY | | |
| 310 | QY | | |
| 311 | QY | | |
| 312 | QY | | |
| 313 | QY | | |
| 314 | QY | | |
| 315 | QY | | |
| 316 | QY | | |
| 317 | QY | | |
| 318 | QY | | |
| 319 | QY | | |
| 320 | QY | | |
| 321 | QY | | |
| 322 | QY | | |
| 323 | QY | | |
| 324 | QY | | |
| 325 | QY | | |
| 326 | QY | | |
| 327 | QY | | |
| 328 | QY | | |
| 329 | QY | | |
| 330 | QY | | |
| 331 | QY | | |
| 332 | QY | | |
| 333 | QY | | |
| 334 | QY | | |
| 335 | QY | | |
| 336 | QY | | |
| 337 | QY | | |
| 338 | QY | | |
| 339 | QY | | |
| 340 | QY | | |
| 341 | QY | | |
| 342 | QY | | |
| 343 | QY | | |
| 344 | QY | | |
| 345 | QY | | |
| 346 | QY | | |
| 347 | QY | | |
| 348 | QY | | |
| 349 | QY | | |
| 350 | QY | | |
| 351 | QY | | |
| 352 | QY | | |
| 353 | QY | | |
| 354 | QY | | |
| 355 | QY | | |
| 356 | QY | | |
| 357 | QY | | |
| 358 | QY | | |
| 359 | QY | | |
| 360 | QY | | |
| 361 | QY | | |
| 362 | QY | | |
| 363 | QY | | |
| 364 | QY | | |
| 365 | QY | | |
| 366 | QY | | |
| 367 | QY | | |
| 368 | QY | | |
| 369 | QY | | |
| 370 | QY | | |
| 371 | QY | | |
| 372 | QY | | |
| 373 | QY | | |
| 374 | QY | | |
| 375 | QY | | |
| 376 | QY | | |
| 377 | QY | | |
| 378 | QY | | |
| 379 | QY | | |
| 380 | QY | | |
| 381 | QY | | |
| 382 | QY | | |
| 383 | QY | | |
| 384 | QY | | |
| 385 | QY | | |
| 386 | QY | | |
| 387 | QY | | |
| 388 | QY | | |

macrophage inflammatory protein-1 alpha receptor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 13-Aug-1999
C:Accession: I49339
R:Gao, J.L.; Murphy, P.M.
J. Biol. Chem. 270, 17494-17501, 1995
A:Title: Cloning and differential tissue-specific expression of three mouse beta
A:Reference number: I49339; MUID:95340546; PMID:7542241
A:Accession: I49339
A:Status: preliminary; translated from CE/EMBL/DBPI

A: Molecule type: DNA
A: Residues: 1-355 <RES>
A/Cross-references: EMBL:U28404; NID:g881547; PIDN:AAA89153.1; PID:g881548
C: Superfamily: vertebrate rhodopsin

Alignment Scores:
Pred. No.: 6,34e-76 Length: 355
Score: 953.00 Matches: 185
Percent Similarity: 69.97% Conservative: 55
Best Local Similarity: 53.94% Mismatches: 88
Query Match: 36.26% Indels: 15
DB: 2 Gaps: 4

US-09-938-703-3 (1-1442) x I49339 (1-355)

| | | | |
|----|------|--|------|
| QY | 270 | GACATCAATATATATACATFCGAGCGCCCTGCCAAAAATCAATGTGAAGCAAAATCGCAGCC | 329 |
| Db | 15 | GIuPheAspTyrGIyAspSerThrProCysGlnLysThrAlaValArgAlaPheGIyAla | 34 |
| QY | 330 | CGCCCTCGGCTCCGCTCTACTCACTGGTGTTCATCTTTGGTTTGGGGCAACATGCTG | 389 |
| Db | 35 | GIyLeuLeuProLeuTyrSerLeuValPheIleIleGIyValValGIyAsnValLeu | 54 |
| QY | 390 | GTCACTCCTCATCTCTGATAACTGCAAGAGCGTGAAGACATGACATCACTACCTGCCTC | 449 |
| Db | 55 | ValIleLeuValLeuMetGlnHisArgLeuGlnSerMetThrSerIleTyrLeuPhe | 74 |
| QY | 450 | AACCTGGCATCTCTCACTGCTTTTCCCTCTTACTGTCCCTCTCTGGGCTCAGTAGCT | 509 |
| Db | 75 | AsnLeuAlaValSerAspLeuValPheLeuPheThrLeuProPheTyrIleAspTyrLys | 94 |
| QY | 510 | GCGCCG--CAGTGGCACTTTGGAAATACAAATGTCTCAACTCTTGACAGGCGCTCTATT | 566 |
| Db | 95 | LeuLysAspAspTrpIlePheGlyAspAlaMetCysLysLeuLeuSerGlyPheTyrTyr | 114 |
| QY | 567 | ATAGCGCTTCTCTGGAACTCTTCTCATCATCTCTCCGCAATGATGATGACTACCTGGCT | 626 |
| Db | 115 | LeuGIyLeuTyrSerGluIlePhePheIleIleLeuLeuThrIleAspTyrLeuAla | 134 |
| QY | 627 | GTGTCCTCATGTGTGTGTCTTAAAGCAGCAGCGGTCACTTTGGGGTGGTGACAACT | 686 |
| Db | 135 | IleValHisAlaValPheAlaLeuArgAlaArgThrValThrLeuGIyIleIleThrSer | 154 |
| QY | 687 | GTGATCACTGGGTGGCTGGCTGTGTGGCTCTCCACGAAATCATCTTACAGATCT | 746 |
| Db | 155 | IleIleThrTrpAlaLeuAlaIleLeuAlaSerMetProAlaLeuTyrPhePheLysAla | 174 |
| QY | 747 | CAAAAAGAGCTTCTCATTACACCTGCAGCTCTCATTTTGCATAC----- | 791 |
| Db | 175 | GlnTrpGluPheThrHisArgThrCysSerProHisPheProTyrLysSerLeuLysGln | 194 |
| QY | 792 | -----ATPAAGATPAGTCATPCTGGGCTGGTCCCGCGCTGGCTGTCT | 833 |
| Db | 195 | TrpLysArgPheGlnAlaLeu-LysLeuAsnLeuLeuGIyIleLeuProLeuLeuVal | 214 |
| QY | 834 | CATGTGTCATCTGCTACTCGGGAATCTPAAAACTCTGCTTCGCTCGTCCGAATGAGAAG | 893 |
| Db | 214 | MetIleIleCysTyrAlaGIyIleIleArgIleLeuLeuArg--ArgProSerGlu | 233 |
| QY | 894 | GAGGCACAGGCTGTGAGGCTTATCTTCACCATCATGATTTATTTCTCTCTCTGGGC | 953 |
| Db | 233 | sLysValLysAlaValArgLeuIlePheAlaIleThrLeuLeuPhePheLeuLeuTrp | 253 |
| QY | 954 | TCCCTACACATTTGCTCTCTCTCCGACACCTTCACGAATTCCTTTGGCTCGTAATATG | 1013 |
| Db | 253 | rProTyrAsnLeuSerValPheValSerAlaPheGlnAspValLeuPheThrAsnGlnC | 273 |
| QY | 1014 | CAGTAGCTCTAACAGGTTGGACCAAGCTATGCAGGTGCAGACGACTCTTGGGATGACGA | 1073 |
| Db | 273 | sGlnGlnSerLysHisLeuAspLeuAlaMetGlnValThrGluValIleAlaTyrThrHi | 293 |
| QY | 1074 | CTGCTGTCATCAACCCCATCATCTATGCTTTCTGCGGGAGAGGTTTCAGAACTACCTCT | 1133 |

[illegible]

[illegible]

Dbb 272 gLeuAlaLeuSerValThrGluThrValAlaPheSerHisCysCysLeuAspProLeuII 292

783 TTTCCATACATTAAAGATAGTCATCTTGGGGCTGGTCCTGCCGCTGCTTGTCAATGGTCAT


```
Db 199 -PheLeuThrLeuLysMetAsnIleSerValLeuValLeuProLeuPheIlePheThrPh 218
QY 843 CTGCTACTCGGAACTCTAAACTCTGCTCGGTGTCGAAATGAGAGAGAGACACAG 902
Db 218 eLeuTyValGlnMetArgLysThrLeu-----ArgPheArgGluGlnArgTyrSe 235
QY 903 GGCTGTGAGGCTTATCTCACCATCATGATTGTTATTCTCTCTCTGGGCTCCCTACAA 962
Db 235 rLeuPheLysLeuValPheAlaIleMetValValPheLeuLeuMetIrpAlaProTyrAS 255
QY 963 CATTCCTCTCTCTGAACACCTTCCAGAACTCTTGGCCTGAAATGTCAGTAGCTC 1022
Db 255 nIleAlaPhePheLeuSerThrPheLysGluHisPheSerLeuSerAspCysLysSerSe 275
QY 1023 TAACAGGTGGACCAAGCTATCCAGGTGACAGAGACTCTTGGGATGACGCACCTGCTGCAT 1082
Db 275 rTyrAsnLeuAspLysSerValHisIleThrLysLeuIleAlaThrThrHisCysCysI 295
QY 1083 CAACCCATCATCTATGCTTGTGCGGAGAGAGTTCAGAACTACTCTTAGTCTTCTT 1142
Db 295 eAsnProLeuLeuTyAlaPheLeuAspGlyThrPheSerLysTyrLeu----- 311
QY 1143 CCAAGACACATTGCCAAACGCTTCTGCAAAATGCTGTCTATTATTCACGACAGAGCTCC 1202
Db 312 -----CysArgCysPheHisLeu---ArgSerAsnThrPr 322
QY 1203 CGAGCGAGCAAGCTCAGTTACACCCGATCCACTGGGAGCAGCAA 1248
Db 322 oLeuGlnProArgGlyGlnSerAlaGlnGlyThrSerArgGluGlu 337
```

Search completed: June 3, 2003, 18:51:41
Job time : 65.6953 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 3, 2003, 18:39:54 ; Search time 23.4701 Seconds
(without alignments)
5096.611 Million cell updates/sec

Title: US-09-938-703-3
Perfect score: 2628
Sequence: 1 GAATTCCTCCACACAGCA.....AGTAGATTAGATCCGAATC 1442

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 225784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-O/cyn2_1/USPTO.spool/US09938703/runat_03062003_161406_22808/app_query.fasta_1.1870
-DB-SwissProt_40 -OPMT-Fastan -SUFFIX-rsp -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS-bits -START=1 -END=1 -MATRIX-blosum62 -TRANS-human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE-pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE-LOCAL
-OUTPMT-pct -NORM-ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER-US09938703.ecgn_1.1_38-brunat_03062003_161406_22808 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NGS_SCORES=0 -WAIT -DSFLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | ID | Description |
|------------|--------|---------------|--------|--------------|---------------------|
| 1 | 1745.5 | 66.4 | 352 | 1 CKR5_HUMAN | P51681 homo sapien |
| 2 | 1739.5 | 66.2 | 352 | 1 CKR5_PANTR | P56440 pan troglod |
| 3 | 1734.5 | 66.0 | 352 | 1 CKR5_PONPY | O97881 pongo pygma |
| 4 | 1729.5 | 65.8 | 352 | 1 CKR5_GORGO | P56439 gorilla gor |
| 5 | 1725.5 | 65.7 | 352 | 1 CKR5_TRAPH | O97879 trachypithe |
| 6 | 1720.5 | 65.5 | 352 | 1 CKR5_PAPHA | P56441 papio hamad |
| 7 | 1720.5 | 65.5 | 352 | 1 CKR5_TRAFR | O97878 trachypithe |
| 8 | 1719.5 | 65.4 | 352 | 1 CKR5_PYGBI | O97880 pygathrix b |
| 9 | 1719.5 | 65.4 | 352 | 1 CKR5_PYGNE | O97882 pygathrix n |
| 10 | 1718.5 | 65.4 | 352 | 1 CKR5_HYLLE | O97883 hylobates l |
| 11 | 1717.5 | 65.4 | 352 | 1 CKR5_MACMU | P79436 macaca mula |
| 12 | 1707.5 | 65.0 | 352 | 1 CKR5_CERTO | O62743 cercopithec |
| 13 | 1694.5 | 64.5 | 352 | 1 CKR5_CERAE | P56493 cercopithec |
| 14 | 1475.5 | 56.1 | 354 | 1 CKR5_RAT | O08556 rattus norv |
| 15 | 1473.5 | 56.1 | 354 | 1 CKR5_MOUSE | P51682 mus muscullu |
| 16 | 1328.5 | 50.6 | 373 | 1 CKR2_MOUSE | P51683 mus muscullu |
| 17 | 1326.5 | 50.5 | 373 | 1 CKR2_RAT | O55193 rattus norv |
| 18 | 1315.5 | 50.1 | 360 | 1 CKR2_MACMU | O18793 macaca mula |

| | | | | | |
|----|--------|------|-----|--------------|---------------------|
| 19 | 1191 | 45.3 | 374 | 1 CKR2_HUMAN | P41597 homo sapien |
| 20 | 1003.5 | 38.2 | 355 | 1 CKR1_HUMAN | P32246 homo sapien |
| 21 | 976 | 37.1 | 359 | 1 CKR3_MOUSE | P51678 mus muscullu |
| 22 | 974 | 37.1 | 359 | 1 CKR3_RAT | O54814 rattus norv |
| 23 | 950 | 36.1 | 355 | 1 CKR1_MOUSE | P51675 mus muscullu |
| 24 | 943.5 | 36.1 | 355 | 1 CKR1_MACMU | P56482 macaca mula |
| 25 | 912 | 34.7 | 355 | 1 CKR3_HUMAN | P51677 homo sapien |
| 26 | 902 | 34.3 | 358 | 1 CKR3_MACMU | P56483 macaca mula |
| 27 | 898.5 | 34.2 | 358 | 1 CKR3_CAVPO | O92213 cavia porce |
| 28 | 891.5 | 33.9 | 360 | 1 CKR4_HUMAN | P51679 homo sapien |
| 29 | 884 | 33.6 | 355 | 1 CKR3_CERAE | P56492 cercopithec |
| 30 | 878.5 | 33.4 | 360 | 1 CKR4_MOUSE | P51680 mus muscullu |
| 31 | 776 | 29.5 | 356 | 1 CKR7_MOUSE | P51676 mus muscullu |
| 32 | 716 | 27.2 | 355 | 1 CKR8_HUMAN | P51685 homo sapien |
| 33 | 709.5 | 27.0 | 356 | 1 CKR8_MOUSE | O97665 macaca mula |
| 34 | 708.5 | 27.0 | 353 | 1 CKR8_MOUSE | P56484 mus muscullu |
| 35 | 688.5 | 26.2 | 354 | 1 C3X1_RAT | P35411 rattus norv |
| 36 | 669.5 | 25.5 | 354 | 1 C3X1_MOUSE | O92089 mus muscullu |
| 37 | 644 | 24.5 | 355 | 1 C3X1_HUMAN | P49238 homo sapien |
| 38 | 582.5 | 22.2 | 357 | 1 CKR9_HUMAN | P51686 homo sapien |
| 39 | 573.5 | 21.8 | 378 | 1 CKR7_HUMAN | P32248 homo sapien |
| 40 | 566.5 | 21.6 | 369 | 1 CKR9_MOUSE | O92077 mus muscullu |
| 41 | 556.5 | 21.2 | 378 | 1 CKR2_MOUSE | P47774 mus muscullu |
| 42 | 556.5 | 21.2 | 384 | 1 CKD6_HUMAN | O00590 homo sapien |
| 43 | 555 | 21.1 | 378 | 1 CKD6_MOUSE | O08707 mus muscullu |
| 44 | 554 | 21.1 | 382 | 1 CKD6_RAT | O09027 rattus norv |
| 45 | 552.5 | 21.0 | 374 | 1 CKR6_HUMAN | P51684 homo sapien |

ALIGNMENTS

RESULT 1
CKR5_HUMAN

| | | | | |
|----|--|---|------|--------|
| ID | CKR5_HUMAN | STANDARD: | PRT: | 352 AA |
| AC | P51681 | O14692; O14693; O14695; O14696; O14697; O14698; O14699; | | |
| AC | O14700; O14701; O14702; O14703; O14704; O14705; O14706; O14707; | | | |
| AC | O14708; O15538; OSUPA4; | | | |
| DT | 01-OCT-1996 (Rel. 34, Created) | | | |
| DT | 01-OCT-1996 (Rel. 34, Last sequence update) | | | |
| DT | 15-OCT-2002 (Rel. 41, Last annotation update) | | | |
| DE | C-C chemokine receptor type 5 (C-C CKR-5) (CCR-5) (CCR5) | | | |
| DE | (HIV-1 fusion co-receptor) (CHEMR13) (CD195 antigen). | | | |
| GN | CCR5 OR CKR5 | | | |
| OS | Homo sapiens (Human). | | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | |
| OC | Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo. | | | |
| OX | NCBI_TaxID=9606; | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RX | MEDLINE-96241590; PubMed-8639485; | | | |
| RA | Samson M., Labbe O., Mollereau C., Vassart G., Parmentier M.; | | | |
| RT | "Molecular cloning and functional expression of a new human | | | |
| RT | CC-chemokine receptor gene." | | | |
| RL | Biochemistry 35:3362-3367(1996). | | | |
| RN | [2] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RX | MEDLINE-96291862; PubMed-8663314; | | | |
| RA | Raport C.J., Gesling J., Schweichart V.L., Gray P.W., Charo I.F.; | | | |
| RT | "Molecular cloning and functional characterization of a novel human | | | |
| RT | CC chemokine receptor (CCR5) for RANTES, MIP-1beta, and MIP-1alpha." | | | |
| RL | J. Biol. Chem. 271:17161-17166(1996). | | | |
| RN | [3] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RX | MEDLINE-96295970; PubMed-8699119; | | | |
| RA | Combadere C., Ahuja S.K., Tiffany H.L., Murphy P.M.; | | | |
| RT | "Cloning and functional expression of CC CKR5, a human monocyte CC | | | |
| RT | chemokine receptor selective for MIP-1(alpha), MIP-1(beta), and | | | |
| RT | RANTES." | | | |
| RL | J. Leukoc. Biol. 60:147-152(1996). | | | |
| RN | [4] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RA | McCombie W.R., Willson R., Chen E., Gibbs R., Zuo L., Johnson D., | | | |
| RA | Nhan M., Parnell L., Dedhia N., Ansari A., Mardis E., Schutz K., | | | |

RA Guoj L., la Bastide M., Kaplan N., Greco T., Touchman J.,
RA Muzny D., Chen C.N., Evans C., Fitzgerald M., See L.H., Tang M.,
RA Porcel B.M., Dragan Y., Giacalone J., Pae A., Powell E.,
RA Solinsky K.A., Desilva U., Diaz-Perez S., Zhou X., Yu Y.,
RA Watanabe M., Doggett N., Garcia D., Sagripanti J.L.,
RL Submitt (MAY-1997) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE-98001387; PubMed-9343222;
RA Kumann S.E., Platt E.J., Kozak S.L., Kabat D.;
RT "Polymorphisms in the CCR5 genes of African green monkeys and mice
RT implicate specific amino acids in infections by simian and human
RT immunodeficiency viruses.";
RL J. Virol. 71:8642-8656(1997).
RN [6]
RP SEQUENCE FROM N.A., AND POLYMORPHISMS.
RX MEDLINE-98022612; PubMed-935954;
RA Zhang L., Carruthers C.D., He T., Huang Y., Cao Y., Wang G., Hahn B.,
RA Ho D.D.;
RT "HIV type 1 subtypes, coreceptor usage, and CCR5 polymorphism.";
RL AIDS Res. Hum. Retroviruses 13:1357-1366(1997).
RN [7]
RP SEQUENCE FROM N.A.
RX MEDLINE-98049523; PubMed-9388201;
RA Mummidi S., Ahuja S.S., McDaniel B.L., Ahuja S.K.;
RT "The human CC chemokine receptor 5 (CCR5) gene. Multiple transcripts
RT with 5'-end heterogeneity, dual promoter usage, and evidence for
RT polymorphisms within the regulatory regions and noncoding exons.";
RL J. Biol. Chem. 272:30662-30671(1997).
RN [8]
RP SEQUENCE FROM N.A., AND VARIANT ARG-178.
RA Magierowska M., Barre-Sinoussi F., Isaifras H., Theodorou I.,
RA Debre P.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
RN [9]
RP CHARACTERIZATION OF ITS HIV-1 RECEPTOR FUNCTION.
RX MEDLINE-96260017; PubMed-8649511;
RA Deng H., Liu R., Ellmeier W., Choe S., Uuttmaz D., Burkhardt M.,
RA di Marzio P., Marmon S., Sutton R.E., Hill C.M., Davis C.B.,
RA Pelper S.C., Schall T.J., Littman D.R., Landau N.R.;
RT "Identification of a major co-receptor for primary isolates of
RT HIV-1.";
RL Nature 381:661-666(1996).
RN [10]
RP CHARACTERIZATION OF ITS HIV-1 RECEPTOR FUNCTION.
RX MEDLINE-96260018; PubMed-8649512;
RA Dragic T., Litwin V., Allaway G.P., Martin S.R., Huang Y.,
RA Nagashima K.A., Cayanan C., Maddon P.J., Koup R.A., Moore J.P.,
RA Paxton W.A.;
RT "HIV-1 entry into CD4+ cells is mediated by the chemokine receptor
RT CC-CKR-5.";
RL Nature 381:667-673(1996).
RN [11]
RP SULFATION.
RX MEDLINE-99189752; PubMed-10089882;
RA Farzan M., Mirzabekov T., Kolchinsky P., Wyatt R., Cayabyab M.,
RA Gerard N.P., Gerard C., Sodroski J., Choe H.;
RT "Tyrosine sulfation of the amino terminus of CCR5 facilitates HIV-1
RT entry.";
RL Cell 96:667-676(1999).
CC -1- FUNCTION: RECEPTOR FOR A C-C TYPE CHEMOKINE. BINDS TO MIP-1-ALPHA,
CC MIP-1-BETA AND RANTES AND SUBSEQUENTLY TRANSDUCES A SIGNAL BY
CC INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL. MAY PLAY A ROLE
CC IN THE CONTROL OF GRANULOCYTIC LINEAGE PROLIFERATION OR
CC DIFFERENTIATION. ACTS AS CO-RECEPTOR WITH CD4 FOR PRIMARY NON-
CC SYNCYTUM-INDUCING STRAINS (NSI) (MACROPHAGE-TROPIC) OF HIV-1
CC VIRUS. IT PROMOTES ENV-MEDIATED FUSION OF THE VIRUS.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: FOUND IN PROMYELOCYTIC CELLS.
CC -1- PTM: SULFATION CONTRIBUTES TO THE EFFICIENCY OF HIV-1 ENTRY.
CC -1- PTM: MODIFIED BY O-LINKED GLYCOSYLATION, BUT NOT BY N-LINKED
CC GLYCOSYLATION.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

CC --- THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X91492; CAA62796.1; -
CC EMBL; U54994; AAC50598.1; -
CC EMBL; U57840; AAB17071.1; -
CC EMBL; U95626; AAB57793.1; -
CC EMBL; U83326; AAC51797.1; -
CC EMBL; AF011500; AAB65700.1; -
CC EMBL; AF011501; AAB65701.1; -
CC EMBL; AF011502; AAB65702.1; -
CC EMBL; AF011503; AAB65703.1; -
CC EMBL; AF011505; AAB65705.1; -
CC EMBL; AF011506; AAB65706.1; -
CC EMBL; AF011507; AAB65707.1; -
CC EMBL; AF011508; AAB65708.1; -
CC EMBL; AF011509; AAB65709.1; -
CC EMBL; AF011510; AAB65710.1; -
CC EMBL; AF011511; AAB65711.1; -
CC EMBL; AF011512; AAB65712.1; -
CC EMBL; AF011513; AAB65713.1; -
CC EMBL; AF011514; AAB65714.1; -
CC EMBL; AF011515; AAB65715.1; -
CC EMBL; AF011516; AAB65716.1; -
CC EMBL; AF011517; AAB65717.1; -
CC EMBL; AF011518; AAB65718.1; -
CC EMBL; AF011519; AAB65719.1; -
CC EMBL; AF011520; AAB65720.1; -
CC EMBL; AF011521; AAB65721.1; -
CC EMBL; AF011522; AAB65722.1; -
CC EMBL; AF011523; AAB65723.1; -
CC EMBL; AF011524; AAB65724.1; -
CC EMBL; AF011525; AAB65725.1; -
CC EMBL; AF011526; AAB65726.1; -
CC EMBL; AF011527; AAB65727.1; -
CC EMBL; AF011528; AAB65728.1; -
CC EMBL; AF011529; AAB65729.1; -
CC EMBL; AF011530; AAB65730.1; -
CC EMBL; AF011531; AAB65731.1; -
CC EMBL; AF011532; AAB65732.1; -
CC EMBL; AF011533; AAB65733.1; -
CC EMBL; AF011534; AAB65734.1; -
CC EMBL; AF011535; AAB65735.1; -
CC EMBL; AF011536; AAB65736.1; -
CC EMBL; AF011537; AAB65737.1; -
CC EMBL; AF011538; AAB65738.1; -
CC EMBL; AF011539; AAB65739.1; -
CC EMBL; AF052539; AAD18131.1; -
CC Genew; HGNC:1606; CCR5.
CC MIM; 601373; -
CC InterPro; IPR000276; GPCR_Rhodpsn.
CC Pfam; PF00001; 7tm1; 1.
CC PRINTS; PR00237; GPCRHOOPS.
CC PROSITE; PS00237; G-PROTEIN RECF_F1_1; 1.
CC PROSITE; PS00262; G-PROTEIN RECF_F1_2; 1.
CC KW G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation;
CC Polymorphism.
CC FT DOMAIN 1 30 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 31 58 1 (POTENTIAL).
CC FT DOMAIN 59 68 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 69 89 2 (POTENTIAL).
CC FT DOMAIN 90 102 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 103 124 3 (POTENTIAL).
CC FT DOMAIN 125 141 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 142 166 4 (POTENTIAL).
CC FT DOMAIN 167 198 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 199 218 5 (POTENTIAL).
CC FT DOMAIN 219 235 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 236 260 6 (POTENTIAL).
 FT DOMAIN 261 277 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 278 301 7 (POTENTIAL).
 FT DOMAIN 302 352 CYTOPLASMIC (POTENTIAL).
 FT DISULFID 101 178 BY SIMILARITY.
 FT MOD_RES 3 3 SULFATION.
 FT MOD_RES 10 10 SULFATION.
 FT MOD_RES 14 14 SULFATION.
 FT MOD_RES 15 15 SULFATION.
 FT VARIANT 10 10 /FTID=VAR_003481.
 FT VARIANT 29 29 Y -> D (IN INCCR5-71A).
 FT VARIANT 31 31 A -> S (IN DBSNP:1800939).
 FT VARIANT 31 31 /FTID=VAR_011839.
 FT VARIANT 31 31 R -> H (IN INCCR5-72A).
 Alignment Scores:
 Pred. No.: 5,25e-131 Length: 352
 Score: 1745.50 Matches: 340
 Percent Similarity: 96.60% Conservative: 1
 Best Local Similarity: 96.32% Mismatches: 0
 Query Match: 66.42% Indels: 12
 DB: 1 Gaps: 1

US-09-938-703-3 (1-1442) x CRR5_HUMAN (1-352)
 QY 240 ATGATTATCAAGTGTCAAGTCCCAATCTATGACATCAATATTATATACATCGAGCCCTGC 299
 Db 1 MetAspTyrGlnValSerSerProIleTyrAspIleAsnTyrTyrThrSerGluProCys 20
 QY 300 CAAAATCAATGTGAAGCAAAATCGCAGCCGCTCTGCTCCGCTCTACTCACTGGTG 359
 Db 21 GlnLysIleAsnValLysGlnIleAlaAlaArgLeuLeuProLeuTyrSerLeuVal 40
 QY 360 TTCACTTTGGTTTGGGCAACATGCTGTCATCTCTATCTGATTAATCTGCAAAAGG 419
 Db 41 PheIlePheGlyPheValGlyAsnMetLeuValIleLeuIleAsnCysLysArg 60
 QY 420 CTGAAGAGCATGACTGACTTACCTGCTCACTGCGCCATCTCTGACCTGTTTTCCTT 479
 Db 61 LeuLysSerMetThrAspIleTyrLeuLeuAsnLeuAlaIleSerAspLeuPheLeu 80
 QY 480 CTACTGTCCTCTGCTGGCTCACTATGCTGCCCGCAGGGGACTTTGGAATACAAATG 539
 Db 81 LeuThrValProPheTyrAlaHisTyrAlaAlaGlnTyrAspPheGlyAsnThrMet 100
 QY 540 TGTCACTCTTCACAGGCTCTATTATAGCTCTCTCTGCGAATCTTTCATCATC 599
 Db 101 CysGlnLeuLeuThrGlyLeuTyrPheIleGlyPheSerGlyLeuPhePheIle 120
 QY 600 CTCCTGACATCGATPAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 659
 Db 121 LeuLeuThrIleAspArgTyrLeuAlaValValHisAlaValPheAlaLeuLysAlaArg 140
 QY 660 ACGGTACCTTTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 719
 Db 141 ThrValThrPheGlyValValThrSerValIleThrTyrValValAlaValPheAlaSer 160
 QY 720 CTCCTGAGGAATCATCTTACAGATCTCAAAAGAGCTCTTCAATACACCTGCGAGCTCT 779
 Db 161 LeuProGlyIleIlePheThrArgSerGlnLysGluGlyLeuHisTyrThrCysSerSer 180
 QY 780 CATTTTCCATAC-----ATTAAAGATAGTGCAT 806
 Db 181 HisPheProTyrSerGlnTyrGlnPheThrLysAsnPheGlnThrLeuLysIleValI 200
 QY 807 CTGGGCGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 866
 Db 200 eLeuGlyLeuValLeuProLeuLeuValMetValIleCysTyrSerGlyIleLeuLysTh 220
 QY 867 TCTCTTGGTGTGCAATGAG 926
 Db 220 fLeuLeuArgCysArgAsnGlnLysLysArgHisArgAlaValArgLeuIlePheThrI 240

QY 927 CAGATTTGTTATTTCCTCTCTGCGCTCCCTACAAATTCCTCTCTCTCTCTCTCTCT 986
 Db 240 eMetIleValTyrPheLeuPheTyrPalaProTyrAsnIleValLeuLeuAsnThrPh 260
 QY 987 CCAGGAATCTTTGGCTGAATAATTCAGTACTTACAGGTTGACCAAGCTATGCA 1046
 Db 260 eGlnGluPheGlyLeuAsnAsnCysSerSerAsnArgLeuAspGlnAlaMetG 280
 QY 1047 GGTGACAGACACTCTTTGGGATGACGACTCTGTCATCAACCCCATCATCTATGCTTTGT 1106
 Db 280 nValThrGluThrLeuGlyMetThrHisCysLysIleAsnProIleIleTyrAlaPheVa 300
 QY 1107 CGGGGAGAAGTTCAGAACTACCTCTTAGTCTTCTTCCAAAGCACATGTCGCAACGCTT 1166
 Db 300 IgLyGluLysPheArgAsnTyrLeuValPhePheGlnLysHisIleAlaLysArgPh 320
 QY 1167 CTGCAAAATGCTGTCTATTTCAGAGAGGCTCCGAGGAGGAGGAGGAGGAGGAGGAG 1226
 Db 320 eCysLysCysCysSerIlePheGlnGlnGlnAlaProGluArgAlaSerSerValTyrTh 340
 QY 1227 CGATCCACTGGGAGCAGGAAATATCTGTGGGCTTG 1263
 Db 340 rArgSerThrGlyGluGlnGluIleSerValGlyLeu 352
 RESULT 2
 CRR5_PANTR STANDARD; PRT: 352 AA.
 ID CRR5_PANTR 002778;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE C-C chemokine receptor type 5 (C-C CKR-5) (CCR-5) (CCR5).
 GN CCR5 OR CXCR5.
 OS Pan troglodytes (Chimpanzee).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
 OX NCBI_TaxID=9598;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97268687; PubMed=9108095;
 RA Edinger A.L., Amedee A., Miller K., Doranz B.J., Endres M.,
 RA Sharron M., Samson M., Lu Z.H., Clements J.E., Murphy-Corb M.,
 RA Pelper S.C., Parmentier M., Broder C.C., Doms R.W.;
 RT "Differential utilization of CCR5 by macrophage and T cell tropic
 RT simian immunodeficiency virus strains.";
 RL Proc. Natl. Acad. Sci. U.S.A. 94:4005-4010(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Zimmerman P.A., Buckler-White A., Alkhatib G.;
 RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98022612; PubMed=9359654;
 RA Zhang L., Carruthers C.D., He T., Huang Y., Cao Y., Wang G., Hahn B.,
 RA Ho D.D.;
 RT "HIV type 1 subtypes, coreceptor usage, and CCR5 polymorphism.";
 RL AIDS Res. Hum. Retroviruses 13:1357-1366(1997).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97426118; PubMed=9282822;
 RA Zacharova V., Zachar V., Goustin A.S.;
 RT "Sequence of chemokine receptor gene CCR5 in chimpanzees, a natural
 RT HIV type 1 host.";
 RL AIDS Res. Hum. Retroviruses 13:1159-1161(1997).
 RN [5]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98090115; PubMed=9430250;
 RA Pretet J.-L., Zerbib A., Girard M., Guillet J.-G., Butor C.;
 RT "Chimpanzee CXCR4 and CCR5 act as coreceptors for HIV type 1.";
 RL AIDS Res. Hum. Retroviruses 13:1583-1587(1997).
 CC -1- FUNCTION: RECEPTOR FOR A C-C TYPE CHEMOKINE. BINDS TO MIP-1-ALPHA,
 CC MIP-1-BETA AND RANTES AND SUBSEQUENTLY TRANSDUCES A SIGNAL BY
 CC INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL. MAY PLAY A ROLE

CC IN THE CONTROL OF GRANULOCYTIC LINEAGE PROLIFERATION OR
 CC DIFFERENTIATION.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF005663; AAB62557.1; -
 DR EMBL; U94329; AAB58446.1; -
 DR EMBL; AF011542; AAB65742.1; -
 DR EMBL; U97666; AAC51670.1; -
 DR EMBL; AF011540; AAB65740.1; -
 DR EMBL; U97977; AAC03717.1; -
 DR InterPro; IPR000276; GPCR_Rhodopsn.
 DR Pfam; PF00001; 7tm1.1; 1.
 DR PRINTS; PR00237; GPCRHOPOPSN.
 DR PROSITE; PS00237; G-PROTEIN_RECP_FL_1; 1.
 DR PROSITE; PS0262; G-PROTEIN_RECP_FL_2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation.
 FT DOMAIN 1 30 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 31 58 1 (POTENTIAL).
 FT DOMAIN 59 68 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 69 89 2 (POTENTIAL).
 FT DOMAIN 90 102 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 103 124 3 (POTENTIAL).
 FT DOMAIN 125 141 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 142 166 4 (POTENTIAL).
 FT DOMAIN 167 198 5 (POTENTIAL).
 FT TRANSMEM 199 218 6 (POTENTIAL).
 FT DOMAIN 219 235 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 236 260 6 (POTENTIAL).
 FT DOMAIN 261 277 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 278 301 7 (POTENTIAL).
 FT DOMAIN 302 352 CYTOPLASMIC (POTENTIAL).
 FT DISULFID 101 178 BY SIMILARITY.
 FT MOD_RES 3 3 SULFATION (BY SIMILARITY).
 FT MOD_RES 10 10 SULFATION (BY SIMILARITY).
 FT MOD_RES 14 14 SULFATION (BY SIMILARITY).
 FT MOD_RES 15 15 SULFATION (BY SIMILARITY).
 FT CARBOHYD 268 268 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CONFLICT 123 123 T -> S (IN REF. 1).
 SQ SEQUENCE 352 AA; 40539 MW; 4A3B698B80FE34C CRC64;

Alignment Scores:
 Pred. No.: 1-57e-130 Length: 352
 Score: 1739.50 Matches: 338
 Percent Similarity: 96.60% Conservative: 3
 Best Local Similarity: 95.75% Mismatches: 0
 Query Match: 66.19% Indels: 12
 DB: 1 Gaps: 1

US-09-938-703-3 (1-1442) x CKR5_PANTR (1-352)

QY 240 ATGGATATCAAGTGTCAAGTCCAACTATGACATCAATATATATACATCGAGCGCCGTC 299
 Db 1 MetaAspTyrGlnValSerProIleTyrAspIleAspTyrTyrThrSerGluProCys 20
 QY 300 CAATAATCAATGTGAGCAATCGAGCGCCGCTCCGCTCGCTACTCATCTGGTG 359
 Db 21 GlnIleLeuValSerGlnIleAlaAlaArgLeuLeuProIleTyrSerLeuVal 40
 QY 360 TTCATCTTTGTTTGTGGGCAACATCTGCTCATCTCCATCTGATGAATGCAAAAGG 419
 Db 41 PheIlePheGlyPheValGlyAsnMetLeuValIleLeuIleLeuLeuLeuLeuVal 60
 QY 420 CTGAGAGACATGACTGACATCTACTCTCTCAACCTGGCATCTCTGACCTGTTTCTCT 479

Db 61 LeuLysSerMetThrAspIleTyrLeuLeuAsnLeuAlaIleSerAspLeuPheLeu 80
 QY 480 CTTACTGTCCCTCTCTGGGCTCACTGTGCGCCAGTGGGACTTTTGAATACAAATG 539
 Db 81 LeuThrValProPheTrpAlaHisTyrAlaAlaAlaGlnTrpAspPheGlyAsnThrMet 100
 QY 540 TGTCAACTCTTGACAGGCTCTATTATATAGGCTTCTCTCTGGAATCTTCATCATC 599
 Db 101 CysGlnLeuLeuThrGlyLeuTyrPheIleGlyPheSerGlyIlePhePheIleIle 120
 QY 600 CTCTGCAATTCATGATAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 659
 Db 121 LeuLeuThrIleAspArgTyrLeuAlaIleValIleAlaValPheAlaLeuLysAlaArg 140
 QY 660 ACGGTCACTTTGGGCTGTGACAGTGTGATCACTTGGGTGGTGGTGGTGGTGGTGT 719
 Db 141 ThrValThrPheGlyValValThrSerValIleThrTrpValValAlaValPheAlaSer 160
 QY 720 CTTCCAGGAATCATCTTACAGATCTCAAAAGAGGCTCTTCACTACCTGCGAGCTCT 779
 Db 161 LeuProGlyIleIlePheThrArgSerGlnLysGluGlyLeuHisTyrThrCysSerSer 180
 QY 780 CATTTTCCATAC-----ATTAAGATAGTCAT 806
 Db 181 HisPheProTyrSerGlnTyrGlnPheTrpLysAsnPheGlnThrLeu-LysIleValIle 200
 QY 807 CTTGGGCTGTGCTGCGCTGCTGTGCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTG 866
 Db 200 eLeuGlyLeuValLeuProLeuValMetValIleCysTyrSerGlyIleLeuLysTh 220
 QY 867 TCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 926
 Db 220 rLeuLeuArgCysArgAsnGluLysLysArgHisArgAlaValArgLeuIlePheThrIle 240
 QY 927 CATGATGTTTATTTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 986
 Db 240 eMetIleValTyrPheLeuPheTrpAlaProTyrAsnIleValLeuLeuLeuAsnThrPh 260
 QY 987 CGAGGAATCTTTGGCTGCTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1046
 Db 260 eGlnGluPhePheGlyLeuAsnAsnCysSerSerSerSerSerSerSerSerSerSer 280
 QY 1047 GGTGACAGAGACTCTTGGGATGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1106
 Db 280 nValThrGlnThrLeuGlyMetThrHisCysCysIleAsnProIleIleTyrAlaPheVa 300
 QY 1107 CGGGAGAGAGTTCAGAACTACTCTTACTGCTTCTTCTTCTTCTTCTTCTTCTTCTT 1166
 Db 300 lGlyLysPheArgAsnTyrLeuValPhePheGlnLysHisIleAlaLysArgPh 320
 QY 1167 CTGCAATCTGTTCTTATTTTCCAGCAAGAGGCTCCGAGGAGGAGGAGGAGGAGGAG 1226
 Db 320 eCysLysCysSerIlePheGlnGlnGlnAlaProGluArgAlaSerValTyrThr 340
 QY 1227 CCGATCCACTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1263
 Db 340 rArgSerThrGlyGluGlnGluIleSerValGlyLeu 352

RESULT 3
 CKR5_PONPY
 ID CKR5_PONPY STANDARD; PRT; 352 AA.
 AC O97881;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE C-C chemokine receptor type 5 (C-CR-5) (CC-CR-5) (CCR-5) (CCR5).
 GN CCR5 OR CCR5.
 OS Pongo pygmaeus (Orangutan).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pongo.
 OX NCBI_TaxID=9600;
 RN [1]
 RP SEQUENCE FROM N.A.

| | | | |
|------------|---|---|--------------|
| QY | 420 | CTGAAGACGATGACTGACATCTACTGCTCAACCTGGCCATCTCTGACCTGTTTTCCTT | 479 |
| DB | 61 | LeuLysSerMetThrAspIleTyrLeuLeuAsnLeuAlaIleSerAspLeuPhePheLeu | 80 |
| QY | 480 | CTTACTGTCCCTTCCTGGGCTCACTATGCTGCCGCCAGTGGGACTTTGGAAATACAATG | 539 |
| DB | 81 | LeuThrValProPheThrAlaHisTyrAlaAlaLagInTrpAspPheGlyAsnThrMet | 100 |
| QY | 540 | TGTCAACTCTGTACAGCGGCTCTATTTTATAGAGCTTCTCTCTGGAATCTCTTCATCATC | 599 |
| DB | 101 | CysGlnLeuLeuThrGlyLeuTyrPheIleGlyPhePheSerGlyIlePhePheIleIle | 120 |
| QY | 600 | CTCCTGACAACTCGATAGTACTCGGCTGTGCTGCCATGCTGTGTGCTTTTAAAGCCAGG | 659 |
| DB | 121 | LeuLeuThrIleAspArgTyrLeuAlaIleValHisAlaValPheAlaLeuLysAlaArg | 140 |
| QY | 660 | ACGGTCACTTGGGCTGGTGCACAAGTGTACATCTGGCTGGCTGGCTGTGTGGCTCT | 719 |
| DB | 141 | ThrValThrPheGlyValValThrSerValIleThrIrpValValAlaValPheAlaSer | 160 |
| QY | 720 | CTCCAGGAATCATCTTTACCATCTCTCAAAAGAGGTCTTCATTACACGTGCAGCTCT | 779 |
| DB | 161 | LeuProGlyIleIlePheThrArgSerGlnLysGluGlyLeuHisTyrThrCysSerSer | 180 |
| QY | 780 | CATTTTCCATAC-----ATTAAGATAGTCAAT | 806 |
| DB | 181 | HisPheProTyrSerGlnTyrGlnPheTrpLysAsnPheGlnThrLeu-LysIleValIle | 200 |
| QY | 807 | CTTGGGGCTGGTCTGCCGCTGTGTTCATGCTATCTGCTACTCGGAAATCTCTAAAC | 866 |
| DB | 200 | eLeuGlyLeuValLeuProLeuLeuValMetValIleCysTyrSerGlyIleLeuLysTh | 220 |
| QY | 867 | TCGTCTCGGCTTCGAAATGAGAGAGAGGACACAGGGCTGTGAGGCTTATCTTCAAT | 926 |
| DB | 220 | rLeuLeuArgCysArgAsnGluLysArgHisArgAlaValArgLeuIlePheThrIle | 240 |
| QY | 927 | CATGATGTTTATTTCTCTCTGGGCTCCCTACAAACATTTGCTCTCTCTGACACCTT | 986 |
| DB | 240 | eMetIleValTyrPheLeuPheThrPalProTyrAsnIleValLeuLeuLeuAsnThrPh | 260 |
| QY | 987 | CCAGGAATCTTTGGCCGTGAATATTCGCTAGTACTCTTAACAGTTGGACCAAGCTATGCA | 1046 |
| DB | 260 | eGlnGlnPhePheGlyLeuAsnAsnCysSerSerAsnArgLeuAspGlnAlaMetG | 280 |
| QY | 1047 | GGTGACAGACACTCTTTGGGATGAGCGACTGCTGTCATCAAGCCCATCATCTATGCTTCT | 1106 |
| DB | 280 | nValThrGluThrLeuGlyMetThrHisCysCysIleAsnProIleTyrAlaPheVa | 300 |
| QY | 1107 | CGGGAGAGTTCAGAACTACTCTTAGCTCTTCTTCCAAAGACACATGTCGAAGCTT | 1166 |
| DB | 300 | lGlyGluIysPheArgAsnTyrLeuLeuValPhePheGlnLysHisIleAlaLysArgPh | 320 |
| QY | 1167 | CTGCAATATGCTGTCTATTTCACAGACAGGCTCCCGAGCGAGCAAGCTCACTTTACAC | 1226 |
| DB | 320 | eCysLysCysSerIlePheGlnGlnAlaProGluArgAlaSerSerValIlyrTh | 340 |
| QY | 1227 | CCGATCCACTGGGGACGAGGAATATCTGTGGGCTTG | 1263 |
| DB | 340 | rArgSerThrGlyGluGlnGluIleSerValGlyLeu | 352 |
| RESULT 5 | | | |
| CKR5_TRAPH | | | |
| ID | CKR5_TRAPH | STANDARD; | PRT; 352 AA. |
| AC | O97879; | | |
| DT | 30-MAY-2000 (Rel. 39, Created) | | |
| DT | 30-MAY-2000 (Rel. 39, Last sequence update) | | |
| DT | 15-JUN-2002 (Rel. 41, Last annotation update) | | |
| DE | C-C chemokine receptor type 5 (C-C CKR-5) (CC-CKR-5) (CCR-5) (CCR5). | | |
| GN | CCR5 ON CCKBR5. | | |
| OS | Trachypithecus playrei (Phayre's leaf monkey). | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | |
| OC | Mammalia; Eutheria; Primates; Catarrhini; Gercopithecidae; Colobinae; | | |

Trachypithecus.
NCBI_TaxID=61618;
[1]
SEQUENCE FROM N.A.
MEDLINE=99416438; PubMed=10486970;
Zhang Y.-W., Zhang O.A., Zhang Y.-P.;
"Sequence evolution of the CCR5 chemokine receptor gene in primates.";
Mol. Biol. Evol. 16:1145-1154(1999).
-1- FUNCTION: RECEPTOR FOR A C-C TYPE CHEMOKINE. BINDS TO MIP-1-ALPHA,
MIP-1-BETA AND RANTES AND SUBSEQUENTLY TRANSDUCES A SIGNAL BY
INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL. MAY PLAY A ROLE
IN THE CONTROL OF GRANULOCYTIC LINEAGE PROLIFERATION OR
DIFFERENTIATION.
-1- SUBCELLULAR LOCATION: Integral membrane protein.
-1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
EMBL: AF075443; AAD19855.1; -
InterPro: IPR000276; GPCR_Rhodpsn.
Pfam: PF00001; 7tm.1.1;
PRINTS: PR00237; GPCRHOODPSN
PROSITE: PS00237; G-PROTEIN-RECEP_FL_1; 1.
PROSITE: PS00262; G-PROTEIN-RECEP_FL_2; 1.
G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation.
FT DOMAIN 1 30 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 31 58 1 (POTENTIAL).
FT DOMAIN 59 68 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 69 89 2 (POTENTIAL).
FT DOMAIN 90 102 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 103 124 3 (POTENTIAL).
FT DOMAIN 125 141 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 142 166 4 (POTENTIAL).
FT DOMAIN 167 198 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 199 218 5 (POTENTIAL).
FT DOMAIN 219 235 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 236 260 6 (POTENTIAL).
FT DOMAIN 261 277 7 (POTENTIAL).
FT TRANSMEM 278 301 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 302 352 BY SIMILARITY.
FT DISULFID 101 178 SULFATION (BY SIMILARITY).
FT MOD_RES 3 3 SULFATION (BY SIMILARITY).
FT MOD_RES 10 10 SULFATION (BY SIMILARITY).
FT MOD_RES 14 14 SULFATION (BY SIMILARITY).
FT MOD_RES 15 15 SULFATION (BY SIMILARITY).
SQ SEQUENCE 352 AA; 40509 MW; 4366FE148D3A5938F CRC64;
Alignment Scores:
Pred. No.: 2.03e-129 Length: 352
Score: 1725.50 Matches: 333
Percent Similarity: 96.32% Conservative: 7
Best Local Similarity: 94.33% Mismatches: 1
Query Match: 65.66% Indels: 12
DB: Gaps: 1
US-09-938-703-3 (1-1442) x CCR5_TRAPH (1-352)
QY 240 ATGGATTATCAAGTCTCAAGTCCAACTATCATCATCAATATTATATACATGGAGCCCTGC 4199
Db 1 MetAspTyrGlnValSerSerProThrTyrAspIleAspTyrTyrThrSerGluProCys 20
QY 300 CAAAATAATCAATGTGAAGCAATATCGACGCCGCCCTCGCTCGCTCTACTACTCATGTG 359
Db 21 GlnLysValAsnValLysGlnIleAlaAlaArgLeuLeuProLeuTyrSerIleuVal 40
QY 360 TTCATCTTTGGTTTGGGCAACATCGCTGTCATCTCTCATCTCTCATTAATACTGCAAAAG 4199

TC-00-029-703-3 (1-1442) x CKR5 TRAPH (1-352)

CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 CC Cercopithecoidea; Papio.
 CC NCBI_TaxID=9557, 9555;
 CC [1]

RP SEQUENCE FROM N.A.
 RC SPECIES-P.hamadryas;
 RX MEDLINE=97268687; PubMed=9108095;
 RA Edinger A.L., Amedee A., Miller K., Doranz B.J., Endres M.,
 RA Sharron M., Samson M., Lu Z.H., Clements J.E., Murphy-Corb M.,
 RA Peiper S.C., Parmentier M., Broder C.C., Doms R.W.;
 RT "Differential utilization of CCR5 by macrophage and T cell tropic
 RT simian immunodeficiency virus strains";
 RL Proc. Natl. Acad. Sci. U.S.A. 94:4005-4010(1997).
 RL [2]

RP SEQUENCE FROM N.A.
 RC SPECIES-P.hamadryas;
 RX MEDLINE=99210133; PubMed=10195758;
 RA Sakkena N.K., Wang B., Novembre F.J., Bolton W., Smit I.K., Lal R.B.;
 RT "Species-specific changes in the CCR5 gene from African and Asian
 RT nonhuman primates";
 RL AIDS Res. Hum. Retroviruses 15:479-483(1999).
 RL [3]

RP SEQUENCE FROM N.A.
 RC SPECIES-P.humanus;
 RC Benton P.A., Timanus D.K., Shearer M.H., Lee D.R., Kennedy R.C.;
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: RECEPTOR FOR A C-C TYPE CHEMOKINE. BINDS TO MIP-1-ALPHA,
 CC MIP-1-BETA AND RANTES AND SUBSEQUENTLY TRANSDUCES A SIGNAL BY
 CC INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL. MAY PLAY A ROLE
 CC IN THE CONTROL OF GRANULOCYTIC LINEAGE PROLIFERATION OR
 CC DIFFERENTIATION.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: AF005658; AB062552.1;
 CC EMBL: AF105287; AAD20556.1;
 CC EMBL: AF105288; AAD20557.1;
 CC EMBL: AF105289; AAD20558.1;
 CC EMBL: AF105290; AAD20559.1;
 CC EMBL: AF023452; AAC63830.1;
 CC InterPro: IPR000276; GPCR_Rhodop.
 CC Pfam: PF00001; 7tm.1.1;
 CC PRINTS: PR00237; GPCRHOOPS.
 CC PROSITE: PS00237; G-PROTEIN RECEPTOR F1.1; 1.
 CC PROSITE: PS00262; G-PROTEIN RECEPTOR F1.2; 1.
 CC G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation.
 CC DOMAIN 1 30 EXTRACELLULAR (POTENTIAL).
 CC TRANSMEM 31 58 1 (POTENTIAL).
 CC DOMAIN 59 68 CYTOPLASMIC (POTENTIAL).
 CC TRANSMEM 69 89 2 (POTENTIAL).
 CC DOMAIN 90 102 EXTRACELLULAR (POTENTIAL).
 CC TRANSMEM 103 124 3 (POTENTIAL).
 CC DOMAIN 125 141 CYTOPLASMIC (POTENTIAL).
 CC TRANSMEM 142 166 4 (POTENTIAL).
 CC DOMAIN 167 198 EXTRACELLULAR (POTENTIAL).
 CC TRANSMEM 199 218 5 (POTENTIAL).
 CC DOMAIN 219 235 CYTOPLASMIC (POTENTIAL).
 CC TRANSMEM 236 260 6 (POTENTIAL).
 CC DOMAIN 261 277 EXTRACELLULAR (POTENTIAL).
 CC TRANSMEM 278 301 7 (POTENTIAL).
 CC DOMAIN 302 352 CYTOPLASMIC (POTENTIAL).
 CC DISULFID 101 178 SULFATION (BY SIMILARITY).
 CC MOD_RES 3 3 SULFATION (BY SIMILARITY).
 CC MOD_RES 10 10 SULFATION (BY SIMILARITY).

FT MOD_RES 14 14 SULEFATION (BY SIMILARITY).
 FT MOD_RES 15 15 SULEFATION (BY SIMILARITY).
 FT CARBOHYD 268 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 352 AA; 40489 MW; 5E15049A9BA1FE8B2 CRC64;
 Alignment Scores:
 Pred. No.: 5,07e-129 Length: 352
 Score: 1720.50 Matches: 333
 Percent Similarity: 96.03% Conservatives: 6
 Best Local Similarity: 94.33% Mismatches: 2
 Query Match: 65.47% Indels: 12
 DB: 1 Gaps: 1

US-09-938-703-3 (1-1442) x CRR5_PAPHA (1-352)

QY 240 ATGGATTATCAAGTGTCAAGTCCCAATCTATGACATCAATATTATATACATGCGGCGCTGC 299
 Db 1 MetaspryGlnValSerSerProThrTyAspIleAspTyThrSerGluProCys 20
 QY 300 CAAATAATCAATGTGAAGCAAAATCGACGCGCTCTCTGCTCGCTCTACTACTACTGCTG 359
 Db 21 GlnLysIleAsnValLysGlnIleAlaAlaArgLeuLeuProLeuTySerLeuVal 40
 QY 360 TTCATCTTGTGTTTGTGGGCAACATGCTGTCATCTCATCTCATCTGATAAATCAATG 419
 Db 41 PheIlePheGlyPheValGlyAsnIleLeuValValLeuIleAsnCysLysArg 60
 QY 420 CTGAAGAGCATGACTGACATCTACTCTCAACCTGGCCATCTCTGACCTGTTTCCTT 479
 Db 61 LeuLysSerMetThrAspIleTyLeuAsnLeuAlaIleSerAspLeuLeuPheLeu 80
 QY 480 CTTACTGTCCCTTCTGGGCTCACTGCTGCGCCAGTGGGACTTGTGAATACATG 539
 Db 81 LeuThrValProPheTrpAlaHisTyAlaAlaGlnTrpAspPheGlyAsnThrMet 100
 QY 540 TGTCACTCTGTGACGGCTCTATTTATAGCTTCTCTCTGGAATCTTCTCATCATC 599
 Db 101 CysGlnLeuLeuThrGlyLeuTyPheIleGlyPhePheSerGlyIlePhePheIle 120
 QY 600 CTCTGCAATCATGATAGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 659
 Db 121 LeuLeuThrIleAspArgTyLeuAlaIleValHisAlaValPheAlaLeuLysAlaArg 140
 QY 660 AGGTCACCTTGGGGTGGTGACAGTGATCATCTGGGTGGTGGTGGTGGTGGTGGTGGTGGT 719
 Db 141 ThrValThrPheGlyValValThrSerValIleThrTrpValValAlaValPheAlaSer 160
 QY 720 CTCCAGGAATCATCTTTACCAAGATCTCAAAAGAGGCTTTCATCTACCTGACCTGCTCT 779
 Db 161 LeuProGlyIleIlePheThrArgSerGlnArgGluGlyLeuHisTyThrCysSerSer 180
 QY 780 CATTTTCCATAC-----ATTAAGATAGTCAT 806
 Db 181 HisPheProTySerGlnTyGlnPheTrpLysAsnPheGlnThrLeu-LysIleValIle 200
 QY 807 CTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 866
 Db 200 LeuGlyLeuValLeuProLeuLeuValMetValIleCysTySerGlyLeuLysTh 220
 QY 867 TCTGCTTGGTGTGCGAATAG 926
 Db 220 rLeuLeuArgCysArgAsnGlnLysLysArgHisArgAlaValArgLeuIlePheThrIle 240
 QY 927 CATGATTTGTTATTTCTCTCTGGGCTCCCTACAAATGCTCTCTCTCTCTCTCTCTCTCTCT 986
 Db 240 eMetIleValTyPheLeuPheTrpAlaProTyAsnIleValLeuLeuLeuAsnThrPh 260
 QY 987 CCAGGAATCTTGGCTGAATATGAGTAGTCTTAACAGGTGGAGGTGGAGGTGGAGGTGGAGGT 1046
 Db 260 eGlnGluPhePheGlyLeuAsnAsnCysSerSerSerAsnArgLeuAspGlnAlaMetG 280
 QY 1047 GGTGACAGAGACTCTTGGGATGACGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1106
 Db 1106

Db 280 nValThrGluThrLeuGlyMetThrHisCysCysAlaLeuProIleIleValAlaPheVa 300
 QY 1107 CGGAGAGAGTTCAGAACTACCTCTTCTTCTTCCAAAGCACATGCGCAAGCGTT 1166
 Db 300 LGlyGluysPheArgasnTrpLeuValPhePheGlnLysHisIleAlaLysArgPh 320
 QY 1167 CTGCAATCTGTTCTTCTATTTCAGCAAGAGCTCCGAGCGAGCAAGCTCAGTTTACAC 1226
 Db 320 eCysLysCysCysSerIlePheGlnGlnGlnAlaProGluArgAlaSerValTyrTh 340
 QY 1227 CCATCCATCGGGAGCAGCAATATCTGTGGCTTG 1263
 Db 340 rArgSerThrGlyGlnGlnLysSerValGlyLeu 352

RESULT 7

CKR5_TRAFR
 ID CKR5_TRAFR STANDARD; PRT; 352 AA.
 AC 097878;
 DT 30-MAY-2000 (Rel. 39, Created)
 DE 30-MAY-2000 (Rel. 39, Last sequence update)
 DE 15-JUN-2002 (Rel. 41, Last annotation update)
 DE C-C chemokine receptor type 5 (C-C CKR-5) (CC-CKR-5) (CCR-5).
 GN CC5 OR CKR5.
 OS Trachypithecus francoisi (Francois' langur) (Indochinese langur).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Colobinae;
 OC Trachypithecus.
 OX NCBI_TaxID=54180;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99416438; PubMed=10486970;
 RA Zhang Y.-W., Ryder O.A., Zhang Y.-P.;
 RL Mol. Biol. Evol. 16:1145-1154(1999).
 CC -1- FUNCTION: RECEPTOR FOR A C-C TYPE CHEMOKINE. BINDS TO MIP-1-ALPHA,
 CC MIP-1-BETA AND RANTES AND SUBSEQUENTLY TRANSDUCES A SIGNAL BY
 CC INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL. MAY PLAY A ROLE
 CC IN THE CONTROL OF GRANULOCYTIC LINEAGE PROLIFERATION OR
 CC DIFFERENTIATION.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/annouoe/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF075442; AAD19854.1; -
 CC InterPro; IPR000276; GPCR_Rhodpsn.
 CC Pfam; PF00001; 7tm.1; 1.
 CC PRINTS; PR00237; GPCRHHODPSN.
 CC PROSITE; PS00237; G-PROTEIN RECEPTOR FL1; 1.
 CC PROSITE; PS0262; G-PROTEIN RECEPTOR FL2; 1.
 CC G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation.
 FT DOMAIN 1 30
 FT TRANSMEM 31 58
 FT DOMAIN 59 68
 FT TRANSMEM 69 89
 FT TRANSMEM 90 102
 FT DOMAIN 103 124
 FT DOMAIN 125 141
 FT TRANSMEM 142 166
 FT DOMAIN 167 198
 FT TRANSMEM 199 218
 FT DOMAIN 219 235
 FT TRANSMEM 236 260
 FT DOMAIN 261 277
 FT TRANSMEM 278 301
 FT DOMAIN 302 352
 FT DISULFID 101 178
 BY SIMILARITY.

FT MOD_RES 3 3 SULFATION (BY SIMILARITY).
 FT MOD_RES 10 10 SULFATION (BY SIMILARITY).
 FT MOD_RES 14 14 SULFATION (BY SIMILARITY).
 FT MOD_RES 15 15 SULFATION (BY SIMILARITY).
 SQ SEQUENCE 352 AA; 40496 MW; 4366F148C25598F CRC64;
 Alignment Scores:
 Pred. No.: 5, 07e-129 Length: 352
 Score: 1720.50 Matches: 332
 Percent Similarity: 96.03% Conservative: 7
 Best Local Similarity: 94.05% Mismatches: 2
 Query Match: 65.47% Indels: 12
 Db: 1 Gaps: 1

US-09-938-703-3 (1-1442) x CKR5_TRAFR (1-352)

QY 240 ATGGATTATCAAGTGTCAAGTCAATGATGACATCAATATTATATACATGCGAGCCCTGC 299
 Db 1 MetAspTyrGlnValSerSerProThrTyrAspIleAspTyrTyrThrSerGluProCys 20
 QY 300 CAAAAAATCAATGTGAAGCAATCGAGCCGCCCTCTGCTCGCTCTACTACTACTGCTG 359
 Db 21 GlnLysValAsnValLysGlnIleAlaAlaArgLeuLeuProLeuTyrSerLeuVal 40
 QY 360 TTCATCTTGGTTTGTGGCAACATGCTGGTCATCTCTCATCTGATAAATCGCAAAAGG 419
 Db 41 PheIlePheGlyPheValGlyAsnIleLeuValValLeuIleLeuIleAsnCysLysArg 60
 QY 420 CTGAAGAGCATGACTGACATCTACCTGCTCAACCTGGCATCTCTGACCTGTTTCTCTT 479
 Db 61 LeuLysSerMetThrAspIleTyrLeuLeuAsnLeuAlaIleSerAspLeuPhePheLeu 80
 QY 480 CTACTCTCCCTCTCTGGCTCACTATGCTGCCGCCAGTGGCAGCTTGGAAATACAAAG 539
 Db 81 LeuThrValProPheThrAlaHisTyrAlaAlaGlnIlePhePhePheLysThrMet 100
 QY 540 TGTCACATCTTACAGAGGCTCTATTATTAGCTTCTCTCTGGAATCTCTTCTCATCATC 599
 Db 101 CysGlnLeuLeuThrGlyLeuTyrPheIleGlyPhePheSerGlyIlePhePheIleIle 120
 QY 600 CTCCTGCAATCGATGAGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 659
 Db 121 LeuLeuThrIleAspArgTyrLeuAlaIleValHisAlaValPheAlaLeuLysAlaArg 140
 QY 660 ACGTCACCTTTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 719
 Db 141 ThrValThrPheGlyValValThrSerValIleThrTrpValValAlaValAlaSer 160
 QY 720 CTCGAGGAATCATCTTACCAGATCTCAAAAGAGCTTCTCATCATCATCATCATCATCAT 779
 Db 161 LeuProGlyIleIlePheThrArgSerGlnArgGluGlyLeuHisTyrThrCysSerSer 180
 QY 780 CATTTCCTATAC-----ATTAAGATATCAT 806
 Db 181 HisPheProTyrSerGlnTyrGlnPheThrPheLysAsnPheGlnThrLeu-LysIleVal 200
 QY 807 CTGGGGCTGGTCTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 866
 Db 200 eLeuGlyLeuValLeuProLeuValMetValIleCysTyrSerGlyIleLeuLysThr 220
 QY 867 TCTGCTTGGTGTGCAAAATGAG 926
 Db 220 rLeuLeuArgCysArgAsnGlnLysLysArgHisArgAlaValArgLeuIlePheThrI 240
 QY 927 CATGATTTCTTATTTTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 986
 Db 240 eMetIleValTyrPheLeuPheThrAlaProTyrAsnIleValLeuLeuLeuAsnThrPh 260
 QY 987 CAGGAATCTTGGCTGAATTAATTCAGTACTTAACAGGTGGACCAAGCATATGCA 1046
 Db 260 eGlnGluPhePheGlyLeuAsnAsnCysSerSerSerSerSerSerSerSerSerSer 280
 QY 1047 GGTGACAGAGACTCTTGGGATGACCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1106

QY 1047 GGTGACAGAGACTCTGGGATGACGACCTGCTGCATCAACCCATCATCTATGCTTGT 1106
 Db 280 nValThrGluThrLeuGlyMetThrHisCysCysIleAsnProIleIleThrAlaPheVa 300
 QY 1107 CGGGAGAAGTTCAGAACTACTCTTAGTCTCTTCCAAAGACACATTCGCAAGGCTT 1166
 Db 300 lGlyGluysPheargAsnThrLeuValPheGlnHisIleAlaLysArgph 320
 QY 1167 CTCGAATCTGTTCTATTTCAGCAAGAGGCTCCGAGCGAGCAAGCTCAGTTTACAC 1226
 Db 320 eCysLysCysCysSerIlePheGlnGlnAlaProGluArgAlaSerSerValThr 340
 QY 1227 CCGATCCACTGGGAGCAGGAGAAATATCTGTGGGCTTG 1263
 Db 340 rArgSerThrGlyGluHisGluIleSerValGlyLeu 352

RESULT 10
 CKR5_HYLLC STANDARD; PRT; 352 AA.
 AC 097883;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE C-C chemokine receptor type 5 (C-C CKR-5) (CC-CKR-5) (CCR-5) (CCR5).
 GN CKR5 OR CMKR5.
 OS Hylobates leucogony (White-cheeked gibbon).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hylobatidae; Hylobates.
 OX NCBI_TaxID=61853;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99416438; PubMed=10486970;
 RA Zhang Y.-W., Ryder O.A., Zhang Y.-P.;
 RT "Sequence evolution of the CCR5 chemokine receptor gene in primates.";
 RL Mol. Biol. Evol. 16:1145-1154(1999).
 CC -1- FUNCTION: RECEPTOR FOR A C-C TYPE CHEMOKINE. BINDS TO MIP-1-ALPHA,
 CC MIP-1-BETA AND RANTES AND SUBSEQUENTLY TRANSDUCES A SIGNAL BY
 CC INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL. MAY PLAY A ROLE
 CC IN THE CONTROL OF GRANULOCYTIC LINEAGE PROLIFERATION OR
 CC DIFFERENTIATION.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; AF075451; AAD19863.1;
 DR InterPro; IPR000276; GPCR_Rhodpsn.
 DR Pfam; PF00001; 7tm_1.1;
 DR PRINTS; PR00237; GPCRHHODPSN.
 DR PROSITE; PS00237; G-PROTEIN RECP_FL1; 1.
 DR PROSITE; PS0262; G-PROTEIN RECP_FL2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation.
 FT DOMAIN 1 30
 FT TRANSMEM 31 58
 FT DOMAIN 59 68
 FT TRANSMEM 69 89
 FT DOMAIN 90 102
 FT TRANSMEM 103 124
 FT DOMAIN 125 141
 FT TRANSMEM 142 166
 FT DOMAIN 167 198
 FT TRANSMEM 199 218
 FT DOMAIN 219 235
 FT TRANSMEM 236 260
 FT DOMAIN 261 277
 FT TRANSMEM 278 301

FT DOMAIN 302 352 CYTOPLASMIC (POTENTIAL).
 FT DISULFID 101 178 BY SIMILARITY.
 FT MOD_RES 3 3 SULFATION (BY SIMILARITY).
 FT MOD_RES 10 10 SULFATION (BY SIMILARITY).
 FT MOD_RES 14 14 SULFATION (BY SIMILARITY).
 SQ SEQUENCE 352 AA; 40445 MW; 4F8B4F344CEB7C91 CRC64;
 Alignment Scores:
 Pred. No.: 7.3e-129 Length: 352
 Score: 1718.50 Matches: 334
 Percent Similarity: 95.75% Conservative: 4
 Best Local Similarity: 94.62% Mismatches: 3
 Query Match: 65.39% Indels: 12
 Db: 1 Gaps: 1
 US-09-938-703-3 (1-1442) x CKR5_HYLLC (1-352)
 QY 240 ATGGATTATCAAGTGTCAAGTCCAAATCTATGACATCAATTATTATACATCGAGCCCTGC 299
 Db 1 MetAspThrGlnValSerSerProThrTyrAspIleAspThrSerGluProCys 20
 QY 300 CAAAAATCAATGTGAGCAAAATCGAGCCCGCTCTCTGCTCGCTCTACTCTACTGGTG 359
 Db 21 GlnLysIleAsnValLysGlnIleAlaAlaArgLeuLeuProLeuTyrSerLeuVal 40
 QY 360 TTCATCTCTGTTTGTGGCAACATCGCTGCTCATCTCTCTGATTAACGCAAGG 419
 Db 41 PheIlePheGlyPheValGlyAsnMetLeuValIleLeuValLeuIleAsnCysLysArg 60
 QY 420 CTGAAGAGCATGACTGACATCTACTGCTCAACCTGGCCATCTCTGACCTGTTTCTT 479
 Db 61 LeuLysSerMetThrAspIleTyrLeuLeuAsnLeuAlaIleSerAspLeuPheLeu 80
 QY 480 CTTACTGTCCCTTCGCGCTCACTATGCTGCGGCCAGTGGGACTTTGGAATACATG 539
 Db 81 LeuThrValProPheThrAlaHisTyrAlaAlaGlnTrpAspPheGlyAsnThrMet 100
 QY 540 TGTCAACTCTGTACAGGGCTCTATTATAGGCTCTCTCTCTGGAATCTCTTCTATCATC 599
 Db 101 CysGlnLeuLeuThrGlyLeuTyrPheIleGlyPheSerGlyIlePhePheIle 120
 QY 600 CTCCTGCAATCATGAGTACGCTGCTGCTCATGCTGTGTGTTGTTTAAAGCCAGG 659
 Db 121 LeuLeuThrIleAspArgTyrLeuAlaIleValHisAlaValPheAlaLeuLysAlaArg 140
 QY 660 AGGTCACCTTTGGGTGTGACAAAGTGTATCACTTGGGTGGTGGCTGTGCTTCTGCT 719
 Db 141 ThrValThrPheGlyValValThrSerValIleThrTrpValValAlaValPheAlaSer 160
 QY 720 CTCACGAAATCATCTTTACAGATCTCAAAAAGAGGTCTTCAATACACCTGCAGCTCT 779
 Db 161 LeuProGlyIleIlePheThrArgSerGlnLysGluGlyLeuHisTyrThrCysSer 180
 QY 780 CATTTTCCATAC-----ATTAAGATAGTCAT 806
 Db 181 HisPheProTyrSerGlnTyrGlnPheThrLysAsnPheGlnThrLeuLysIleVal 200
 QY 807 CTGGGGCTGTCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 866
 Db 200 eLeuGlyLeuValLeuProLeuValMetValIleCysTyrSerGlyIleLeuLys 220
 QY 867 TCTGCTTGGTGTGCAAAATGAGAAGAGGACAGGGCTGTGAGGCTTATCTTCCAT 925
 Db 220 rLeuLeuArgCysArgAsnGluLysLysArgHisArgAlaValArgLeuIlePheThr 240
 QY 927 CATGATCTTTTATTTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 986
 Db 240 eMetIleValTyrPheLeuPheThrAlaProTyrAsnIleValLeuLeuLeuAsnThrPh 260
 QY 987 CAGGAATCTTGGCTGAAATTAATTGAGTAGTCTTACAGGTGTCAGGCTGACCACTATGCA 1046
 Db 260 eGlnGluPhePheGlyLeuAsnAsnCysSerSerSerAsnArgLeuAspGlnAlaMetG 280

Db 41 PheIlePheGlyPheValGlyAsnIleLeuValValLeuIleAsnCysLysArg 60

QY 420 CTGAAGACGATGACTGACATCTACCTGCTCAACCTGGCCATCTCTGACCTGTTTTCCCTT 479
|||||

Db 61 LeuLysSerMetThrAspIleTyrLeuLeuAsnLeuAlaIleSerAspLeuLeuPheLeu 80

QY 240 ATGGATTATCAAGTCTCAAGTCAATCTATGACATCAATATATATATACATCGAGCCCTGC 299
 Db 1 MetAspTyrGlnValSerSerProThrTyrAspIleAspTyrTyrThrSerGluProCys 20
 QY 300 CAAAAATCAATGTGAAGAAATCGCAGCCGCTCTCTGCTCGCTCTACTCTACTCTGGTG 359
 Db 21 GlnIysIleAsnValLysGlnIleAlaAlaArgLeuLeuProLeuTyrSerLeuVal 40
 QY 360 TTCATCTTGGTTTGGGGCAACATCGTGGTCACTCTCATCTCATCTGATTAATGCAAAAGG 419
 Db 41 PheIlePheGlyPheValGlyAsnIleLeuValValLeuIleLeuIleAsnCysLysArg 60
 QY 420 CTGAAGAGCATGACTGACATCTACCTGCTCAACTGGCCATCTCTGACCTGTTTCCCT 479
 Db 61 LeuIysSerMetThrAspIleTyrLeuLeuAsnLeuAlaIleSerAspLeuPheLeu 80
 QY 480 CTTACTGTCCTTCTGGGCTCACTATGCTGCGCCGAGTGGGACTTTGGAATPACAATG 539
 Db 81 LeuThrValProPheThrAlaHisTyrAlaAlaGlnTrpAspPheGlyAsnThrMet 100
 QY 540 TGTCAACTCTTGACAGGCTCTATTTATAGGCTTCTCTGGAATCTCTTCTCATCATC 599
 Db 101 CysGlnLeuLeuThrGlyLeuTyrPheIleGlyPheSerGlyIlePhePheIle 120
 QY 600 CTCCTGACATCATGATGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 659
 Db 121 LeuLeuThrIleAspTyrLeuAlaIleValHisAlaValPheAlaLeuLysArg 140
 QY 660 ACGTCACTTGGGGTGGTGAACAGTGTGATGATGATGATGATGATGATGATGATGATG 719
 Db 141 ThrValThrPheGlyValValThrSerValIleThrTrpValValAlaValPheAlaSer 160
 QY 720 CTCCTGAGATCATCTTACCATGCTCAAAAGAGTCTTCAATACAGCTGAGCTCT 779
 Db 161 LeuProGlyIleIlePheThrArgSerGlnArgGluGlyLeuHisTyrThrCysSerPro 180
 QY 780 CATTTTCCATAC-----ATTAAAGATAGTCAT 806
 Db 181 HisPheProTyrSerGlnTyrGlnPheTyrLysAsnPheGlnThrLeuLysIleVal 200
 QY 807 TTTGGGCTGGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 866
 Db 200 eLeuGlyLeuValLeuProLeuLeuValMetValIleCysTyrSerGlyLeuLysTh 220
 QY 867 TCTCTCTGCTGCTCAATGAGAGAGAGGACAGGCTGCTGAGCTTATCTTCACCAT 926
 Db 220 rLeuLeuArgCysArgGlnGluLysArgHisArgAlaValArgLeuIlePheThr 240
 QY 927 CATGATTGTTTATTTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 986
 Db 240 eMetIleValTyrPheLeuPheThrPalaProTyrAsnIleValLeuLeuAsnThrPh 260
 QY 987 CCAGGAATCTTGGCTGCTGAATATTCAGTAGCTCTAACAGGTGGACCAAGCTATGCA 1046
 Db 260 eGlnGluPheGlyLeuAsnAsnCysSerSerSerAsnArgLeuAspGlnAlaMetG 280
 QY 1047 GGTGACAGACATCTTGGATGACGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1106
 Db 280 nValThrGluThrLeuGlyMetThrHisCysCysIleAsnProIleIleTyrAlaPheVa 300
 QY 1107 CGGGAGAGTCTGAGAACTACCTCTTGTAGTCTTCTTCAAGACATTCGCAACGCTT 1166
 Db 300 lGlyGluLysPheArgAsnTyrLeuLeuValPhePheGlnIlyHisIleAlaLysArgPh 320
 QY 1167 CTGCAATCTGCTCTATTTTCCAGAGAGGCTCCCGAGGACGAGCTAGTTTACAT 1226
 Db 320 eCysLysCysSerIlePheGlnGlnGluAlaSerGluArgAlaSerSerValTyrTh 340
 QY 1227 CGATTCACCTGGGGAGAGAAATATCTGTGGGCTTG 1263
 Db 340 rArgSerThrGlyGluGlnGluIleSerValGlyLeu 352

RESULT 13

CKR5_CERAE
 ID CKR5_CERAE STANDARD; PRT; 352 AA.
 AC PS6493;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE C-C chemokine receptor type 5 (C-C CKR-5) (CCR-5) (CCR5).
 GN CCR5 OR CCR5.
 OS Cercopithecus aethiops (Green monkey) (Grivet).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
 OC Cercopithecinae; Cercopithecus.
 OX NCBI_TaxID=9534;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RX MEDLINE=98001387; PubMed=9343222;
 RA Kuhlmann S.E., Platt E.J., Kozak S.L., Kabat D.;
 RT "Polymorphisms in the CCR5 genes of African green monkeys and mice
 RT implicate specific amino acids in infections by simian and human
 RT immunodeficiency viruses.";
 RL J. Virol. 71:8642-8656(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Murayama Y., Matsunaga S., Inoue-Murayama M.;
 RT "cDNA sequence of African green monkey CCR-5 chemokine receptor
 RT gene.";
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
 CC !- FUNCTION: RECEPTOR FOR A C-C TYPE CHEMOKINE. BINDS TO MIP-1-ALPHA,
 CC MIP-1-BETA AND RANTES AND SUBSEQUENTLY TRANSDUCES A SIGNAL BY
 CC INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL. MAY PLAY A ROLE
 CC IN THE CONTROL OF GRANULOCYTIC LINEAGE PROLIFERATION OR
 CC DIFFERENTIATION
 CC !- SUBCELLULAR LOCATION: Integral membrane protein.
 CC !- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; U83324; AAC51795.1; -;
 DR EMBL; U83325; AAC51796.1; -;
 DR EMBL; AB015944; BAA31328.1; -;
 DR InterPro; IPR000276; GPCR_Rhodpsn.
 DR Pfam; PF00001; 7tm1.1;
 DR PRINTS; PR00237; GPCRHHODPSN.
 DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
 DR PROSITE; PS00262; G_PROTEIN_RECEP_F1_2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation;
 KW Polymorphism.
 FT DOMAIN 1 30 EXTRACELLULAR (POTENTIAL).
 FT TRANSSEM 31 58 1 (POTENTIAL).
 FT DOMAIN 59 68 CYTOPLASMIC (POTENTIAL).
 FT TRANSSEM 69 89 2 (POTENTIAL).
 FT DOMAIN 90 102 EXTRACELLULAR (POTENTIAL).
 FT TRANSSEM 103 124 3 (POTENTIAL).
 FT DOMAIN 125 141 CYTOPLASMIC (POTENTIAL).
 FT TRANSSEM 142 166 4 (POTENTIAL).
 FT DOMAIN 167 198 EXTRACELLULAR (POTENTIAL).
 FT TRANSSEM 199 218 5 (POTENTIAL).
 FT DOMAIN 219 235 CYTOPLASMIC (POTENTIAL).
 FT TRANSSEM 236 260 6 (POTENTIAL).
 FT DOMAIN 261 277 EXTRACELLULAR (POTENTIAL).
 FT TRANSSEM 278 301 7 (POTENTIAL).
 FT DOMAIN 302 352 CYTOPLASMIC (POTENTIAL).
 FT DISULFD 101 178 BY SIMILARITY.
 FT MOD_RES 3 3 SULFATION (BY SIMILARITY).
 FT MOD_RES 10 10 SULFATION (BY SIMILARITY).
 FT MOD_RES 15 15 SULFATION (BY SIMILARITY).

| | | | |
|-----------|--|---|--------------|
| QY | 1107 | CGGGGAGAGTTCAGAAACTACCTCTTACTCTCTCCAAAGACACATGCGAAAGCGTT | 1166 |
| Db | 300 | LGyLgLuysPheArGAsTyLeuLeuValPhePheGlnLyShISleAlaLyArghP | 320 |
| QY | 1167 | CTGCAAAATGCTGTCTATTTCAGCAGCAAGCGCTCCGCGAGCAGCAAGCTCAGCTTTACAC | 1226 |
| Db | 320 | eCysLyScysSerIlePheGlnGlnGluAlaProGluArgAlaSerSerValIYrTh | 340 |
| QY | 1227 | CCGATCCACTGGGGGAGCAGGAATAATCTGTGGGC | 1260 |
| Db | 340 | rArGSerThrGlyGlnGlnGluThrSerValGly | 351 |
| RESULT 14 | | | |
| ID | CKR5_RAT | STANDARD; | PRT; 354 AA. |
| AC | 008556; | | |
| DT | 01-NOV-1997 (Rel. 35, Created) | | |
| DT | 01-NOV-1997 (Rel. 35, Last sequence update) | | |
| DT | 30-MAY-2000 (Rel. 33, Last annotation update) | | |
| DE | C-C chemokine receptor type 5 (C-C CKR-5) (CCR-5) (MIP-1 alpha receptor). | | |
| DE | CCR5 OR CMKBR5. | | |
| GN | Rattus norvegicus (Rat). | | |
| OS | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | |
| OC | Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus. | | |
| OX | NCBI_TaxID=10116; | | |
| RN | [1] | | |
| RP | SEQUENCE FROM N.A. | | |
| RC | STRAIN=Wistar; TISSUE=Brain; | | |
| RA | MEDLINE=98318173; PubMed=9670989; | | |
| RX | Spieiss O., Gournala N., Boddeke H.W.G.M., Sauter A., Fiebich B.L., | | |
| RA | Berger M., Gebicke-Haerter P.J.; | | |
| RT | "Cloning of rat hiv-1-chemokine coreceptor CKR5 from microglia and | | |
| RT | upregulation of its mRNA in ischemic and endotoxemic rat brain.;" | | |
| RL | J. Neurosci. Res. 53:16-28(1998). | | |
| RL | [2] | | |
| RP | SEQUENCE FROM N.A. | | |
| RC | STRAIN=Sprague-Dawley; | | |
| RA | MEDLINE=98318173; PubMed=9655467; | | |
| RX | Jiang Y., Salafranca M.N., Adhikari S., Xia Y., Feng L., Sonntag M.K., | | |
| RA | Defebvre C.M., Pennell N.A., Streit W.J., Harrison J.K.; | | |
| RT | "Chemokine receptor expression in cultured glia and rat experimental | | |
| RT | allergic encephalomyelitis.;" | | |
| RL | J. Neuroimmunol. 86:1-12(1998). | | |
| CC | -1- FUNCTION: RECEPTOR FOR A C-C TYPE CHEMOKINE. BINDS TO MIP-1-ALPHA, | | |
| CC | MIP-1-BETA AND RANTES AND SUBSEQUENTLY TRANSDUCES A SIGNAL BY | | |
| CC | INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL. | | |
| CC | -1- SUBCELLULAR LOCATION: Integral membrane protein. | | |
| CC | -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS. | | |
| CC | ----- | | |
| CC | This SWISS-PROT entry is copyright. It is produced through a collaboration | | |
| CC | between the Swiss Institute of Bioinformatics and the EMBL outstation - | | |
| CC | the European Bioinformatics Institute. There are no restrictions on its | | |
| CC | use by non-profit institutions as long as its content is in no way | | |
| CC | modified and this statement is not removed. Usage by and for commercial | | |
| CC | entities requires a license agreement (See http://www.isb-sib.ch/announce/ | | |
| CC | or send an email to license@sib-sib.ch). | | |
| CC | ----- | | |
| DR | EMBL; Y12009; CAA72737.1; -. | | |
| DR | EMBL; U77350; AAC03243.1; -. | | |
| DR | InterPro; IPR00276; GPCR_Rhodpsn. | | |
| DR | Pfam; PF00001; 7tm1.1; 1. | | |
| DR | PRINTS; PR00237; GPCRHHODPSN. | | |
| DR | PROSITE; PS00237; G_PROTEIN_RECP_FL_1; 1. | | |
| DR | PROSITE; PS0262; G_PROTEIN_RECP_FL_2; 1. | | |
| KW | G-protein coupled receptor; Transmembrane; Glycoprotein. | | |
| FT | DOMAIN 1 32 EXTRACELLULAR (POTENTIAL). | | |
| FT | TRANSMEM 33 60 1 (POTENTIAL). | | |
| FT | DOMAIN 61 70 CYTOPLASMIC (POTENTIAL). | | |
| FT | DOMAIN 71 91 2 (POTENTIAL). | | |
| FT | DOMAIN 92 104 EXTRACELLULAR (POTENTIAL). | | |
| FT | TRANSMEM 105 126 3 (POTENTIAL). | | |

FT DOMAIN 127 143 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 144 168 4 (POTENTIAL).
 FT DOMAIN 169 200 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 201 220 5 (POTENTIAL).
 FT DOMAIN 221 237 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 238 262 6 (POTENTIAL).
 FT DOMAIN 263 279 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 280 303 7 (POTENTIAL).
 FT DOMAIN 304 354 CYTOPLASMIC (POTENTIAL).
 FT DISULFID 103 180 BY SIMILARITY.
 FT CARBOHYD 270 270 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 354 AA; 41030 MW; 77EDB368AA4C868D CRC64;

Alignment Scores:
 Pred. No.: 1,42e-109 Length: 354
 Score: 1475.50 Matches: 283
 Percent Similarity: 86.48% Conservative: 34
 Best Local Similarity: 79.72% Mismatches: 24
 Query Match: 56.15% Indels: 14
 DB: 1 Gaps: 2

US-09-938-703-3 (1-1442) x CKR5_RAT (1-354)
 QY 240 ATGATTATCAAGTGTCTAGTCCA-----ATCTATGACATCAATTATTATATACATCGGAG 293
 Db 1 MetasphegInGlySerLeuProThrTyrIleTyrAspIleAspTyrSerMetSerAla 20
 QY 294 CCCTGCCAAAATAATGATGAGCAAAATGCGCCGCCCTCCGCTCCGCTCTACTCA 353
 Db 21 ProCysGlnLysValAsnValLysGlnIleAlaAlaGlnLeuLeuProLeuTyrSer 40
 QY 354 CTGGTGTCTATCTTGGTGTGGCAACATGCTGGTCACTCTCTCATCTCTGATAACTGC 413
 Db 41 LeuValPheIlePheGlyPheValGlyAsnMetValPheLeuIleLeuIleSerCys 60
 QY 414 AAAAGGTGAAGAGCATGACTGACATCTACCTGCTCAACCTGGCCATCTCTGACCTGTTT 473
 Db 61 LysLysLeuLysSerMetThrAspIleTyrLeuPheAsnLeuAlaIleSerAspLeuLeu 80
 QY 474 TTCCTCTTACTGTCCTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 533
 Db 81 PheLeuLeuThrLeuProPheTrpAlaHisTyrAlaAlaGlnLeuValPheGlyAsn 100
 QY 534 ACAATGTCTCAACTCTGACAGGCTCTATTATAGCTTCTCTCTGCTGCTGCTGCTGCTGCT 593
 Db 101 IleMetCysLysLeuPheThrGlyIleTyrHisIleGlyTyrPheGlyGlyIlePhePhe 120
 QY 594 ATCATCTCTGACATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 653
 Db 121 IleIleLeuLeuThrIleAspArgTyrLeuAlaIleValHisAlaValPheAlaIleLys 140
 QY 654 GCAGGAGGTGACCTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 713
 Db 141 AlaArgThrValAsnPheGlyValIleThrSerValValThrIrpValAlaValAlaValPhe 160
 QY 714 GCGTCTCTCCAGGATCATCTTTACCATCTCTCAAAAGAGGTCTTCATTACACCTGTC 773
 Db 161 ValSerLeuProGluIleIlePheMetArgSerGlnLysGluGlySerHisTyrThrCys 180
 QY 774 AGC-----TCTCATTTTCCATACATTAAGAT 800
 Db 181 SerProHisPheLeuHisIleGlnTyrArgPheTrpLysHis-PheGluThrLeuLysMe 200
 QY 801 AGTCATCTTGGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 860
 Db 200 ValIleLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 220
 QY 861 AAAAATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 920
 Db 220 uAsnThrLeuPheArgCysArgAsnGluLysLysArgHisArgAlaValArgLeuIlePhe 240
 QY 921 CACCATCATGATTGTTTATTTCTCTTCTGGGCTCCCTACACATGCTCTCTCTCTCTGAA 980

Db 240 eAlaIleMetIleValTyrPheLeuPheTrpThrProTyrAsnIleValLeuLeuTh 260
 QY 981 CACCTTCCAGGAATCTTTGGCTGAATTAATTCAGTAGCTCTTAACAGGTGGACCAAGC 1040
 Db 260 rThrPheGlnGluTyrPheGlyLeuAsnAsnCysSerSerSerAsnArgLeuAspGlnAl 280
 QY 1041 TATGAGGTGACAGACACTCTGGATGACGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1100
 Db 280 aMetGlnValThrGlnThrLeuGlyMetThrHisCysLysLeuAsnProValIleTyrAl 300
 QY 1101 CTTTGTGCGGGAGAGTTCAGAACTACCTTGTAGTCTCTTCCAAAAGACATTCGCAA 1160
 Db 300 aPheValGlyGluLysPheArgAsnTyrLeuSerValPhePheArgLysHisIleVal 320
 QY 1161 ACCTTCTGCAAAATCGTGTCTATTTCCAGCAAGAGGCTCCGAGCGACCAATTCAGT 1220
 Db 320 sArgPheCysLysHisCysSerIlePheGlnGlnValAsnProAspArgValSerSerVa 340
 QY 1221 TTACACCCGATCCACTGGGAGCAGGAATATCTGTGGGCTTG 1263
 Db 340 lTyrThrArgSerThrGlyGlnGluValSerThrGlyLeu 354

RESULT 15
 CKR5_MOUSE
 ID CKR5_MOUSE STANDARD; PRT: 354 AA.
 AC P51682; 061867; P97405; 035313; P97308; 035891;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE C-C chemokine receptor type 5 (C-C CKR-5) (CC-CKR-5) (CCR-5) (MIP-1
 alpha receptor).
 GN CKR5 OR CKR5.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN=129/SVJ; TISSUE=Spleen;
 RX MEDLINE=96205938; PubMed=8631787;
 RA Boring L., Gosling J., Monteclaro F.S., Lusis A.J., Tsou C.-L.,
 Charo I.F.;
 RT "Molecular cloning and functional expression of murine JE (monocyte
 chemoattractant protein 1) and murine macrophage inflammatory protein
 1alpha receptors: evidence for two closely linked C-C chemokine
 receptors on chromosome 9.";
 RL J. Biol. Chem. 271:7551-7558(1996).
 RN [2]
 RC SEQUENCE FROM N.A.
 RP STRAIN=C57BL/6 X CBA; TISSUE=Thymus;
 RX MEDLINE=96278910; PubMed=8662890;
 RA Meyer A., Coyle A.J., Proudfoot A.E.I., Wells T.N.C., Power C.A.;
 RT "Cloning and characterization of a novel murine macrophage
 inflammatory protein-1 alpha receptor.";
 RL J. Biol. Chem. 271:14445-14451(1996).
 RN [3]
 RC SEQUENCE FROM N.A.
 RP STRAIN=129/Ola.
 RX MEDLINE=96278910; PubMed=8662890;
 RA Kuziel W.A., Beck M.A., Dawson T.C., Maeda N.;
 RT Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RC SEQUENCE FROM N.A.
 RP STRAIN=C57BL/6 and NIH Swiss; TISSUE=Liver, Kidney, and Spleen;
 RX MEDLINE=98001387; PubMed=9343222;
 RA Kuhnmann S.E., Platt E.J., Kozak S.L., Kabat D.;
 RT "Polymorphisms in the CCR5 genes of African green monkeys and mice
 implicate specific amino acids in infections by simian and human
 immunodeficiency viruses.";
 RL J. Virol. 71:8642-8656(1997).
 RN [5]
 RC SEQUENCE FROM N.A.
 RP STRAIN=129;
 RX MEDLINE=97404635; PubMed=9261347;

| Alignment Scores: | | | |
|---|-----------|--|-----|
| Pred. No.: | 2.05e-109 | Length: | 354 |
| Score: | 1473.50 | Matches: | 281 |
| Percent Similarity: | 87.04% | Conservative: | 32 |
| Best Local Similarity: | 79.15% | Mismatches: | 28 |
| Query Match: | 56.07% | Indels: | 14 |
| DB: | 1 | Gaps: | 2 |
| US-09-938-703-3 (1-1442) x CKRS_MOUSE (1-354) | | | |
| QY | 240 | ATGATATTACAGTGTCCAGTCCA-----ATCTATGACATCAATTATTATATACATCGGAG | 293 |
| Db | 1 | MetaspheGlnGlySerValProThrTyrlleTyraPilleAspyrGlyMetSerAla | 20 |
| QY | 294 | CCCTGCCAAAAAATCAATGTGAAGCAATCGACGCCGCCCTCGCTCGCTCTACTCA | 353 |
| Db | 21 | ProCysGlnLysIleAsnValLysGlnIleAlaGlnLeuLeuProLeuTySer | 40 |
| QY | 354 | CTGGTCTCATCTTTGGTTTTTGTGGCAACATCGTGGTCATCTCATCTGTATAAATGC | 413 |
| Db | 41 | LeuValPheIlePheGlyPheValGlyAsnMetMetValPheLeuIleLeuIleSerCys | 60 |
| QY | 414 | AAAAGCTGAAGACGACTGACATCTACTCTGCTCAACTGGCCATCTCTGACCTGTTT | 473 |
| Db | 61 | LysLysLeuLysSerValThrAspIleTyLeuLeuAsnLeuAlaIleSerAspLeuLeu | 80 |
| QY | 474 | TTCTCTTCTACTGTCCCTTCTGGGCTCACTATGCTGCGGCCAGTGGGACATTTGGAAT | 533 |
| Db | 81 | PheLeuLeuThrLeuProPheTrpAlaHisTyrAlaAlaAsnGluTrpIlePheGlyAsn | 100 |
| QY | 534 | ACAAATGTGCAACTTTGACAGGGCTCATTTATAGGCTTCTTCTCGGAATCTCTTC | 593 |
| Db | 101 | IleMetCysLysValPheThrGlyValTyrHisIleGlyTyrPheGlyGlyIlePhePhe | 120 |
| QY | 594 | ATCATCTCTCTGACAAFCGATAGGTACCTGGCTGTGCTGCATGCTGTGTTCCTTAAA | 653 |
| Db | 121 | IleIleLeuLeuThrIleAspArgTyrLeuAlaIleValHisAlaValPheAlaLeuLys | 140 |
| QY | 654 | GCACAGAGGGTCACCTTTGGGTGTGACAGTGTCAATCACTTGGGTGGTGGCTGTGTT | 713 |
| Db | 141 | ValArgThrValAsnPheGlyValIleThrSerValValThrTrpValValAlaValPhe | 160 |
| QY | 714 | GCCTCTCTCCAGGAATCAFTTTACACAGATCTCAAAAAGAGTCTTCAATACACCTGC | 773 |
| Db | 161 | AlaSerLeuProGluIleIlePheThrArgSerGlnLysGluGlyPheHisTyrThrCys | 180 |
| QY | 774 | AGCTCTCATTTTCCATAC-----ATTAAGAT | 800 |
| Db | 181 | SerProHisPheProHisThrGlnTyrHisPheTrpLysSerPheGlnThrLeu-LysMe | 200 |
| QY | 801 | AGTCATCTTGGGCTGTCTCGCTCGCTGTGTCATGTCATCTGCTACTCGGGAATCCT | 860 |
| Db | 200 | tValIleLeuSerLeuIleLeuProLeuLeuValMetIleIleCysTyrSerGlyIleLe | 220 |
| QY | 861 | AAAAACTCTGCTTCGGTGTGCAAAATGACAGAAGAGGCACAGGCTGTGAGGCTTATCTT | 920 |
| Db | 220 | uHisThrLeuPheArgCysArgAsnGluLysLysArgHisArgAlaValArgLeuIlePh | 240 |
| QY | 921 | CACCATCATGATTTGTTTATTCTTCTTGGGTCCCTACACATTTGCTCTCTCTGAA | 980 |
| Db | 240 | eAlaIleMetIleValTyrPheLeuPheTrpThrProTyrAsnIleValLeuLeuLeuth | 260 |
| QY | 981 | CACCTTCCAGGAATCTTTGGCTCAATATTCGATAGTCTTAACAGTGTGGACAGC | 104 |
| Db | 260 | rThrPheGlnGluPhePheGlyLeuAsnAsnCysSerSerSerAsnArgLeuAspGlnAl | 280 |
| QY | 1041 | TATCAGGTGTACAGAGACTTTGGGATGAGCGACTGTCATCAACCCCATCATCTATGC | 110 |
| Db | 280 | aMetGlnAlaThrGluThrLeuGlyMetThrHisCysCysLeuAsnProValIleTyrAl | 300 |
| QY | 1101 | CTTTGTGGGGAGAGTTCAAGAACTACCTTCTAGTCTTCTTCCAAAGCACATTTGCCA | 116 |

Db 300 aPheValGlyGluLysPheArgSerTyrLeuSerValPhePheArgLysHisIleVally 320
QY 1161 ACGTTCTGCAAAATGCTTCTATTTCAGCAAGAGGCTCCGAGCAGCAGCTCAGT 1220
|||||
Db 320 sArgPheCysLysArgCysSerIlePheGlnGlnAspAsnProAspArgValSerSerVa 340
|||||
QY 1221 TTACACCGATCCACTGGGGAGCAGCAATATCTGTGGGCTTG 1263
|||||
Db 340 lTyThrArgSerThrGlyGluHisGlnValSerThrGlyLeu 354
|||||

Search completed: June 3, 2003, 18:46:01
Job time : 39.4701 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_p2p model

Run on: June 3, 2003, 18:41:24 ; Search time 103.268 Seconds
(without alignments)
5754.334 Million cell updates/sec

Title: US-09-938-703-3

Perfect score: 2628

Sequence: 1 GAATTCCTCCACAGACCA.....AGTAGATTAGATCCGAATTC 1442

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 1343160

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-DB=/cygn2_1/USPTO_spool/US09938703/runat_03062003_161407_22827/app_query.fasta.1.1870
-Q=/cygn2_1/USPTO_spool/US09938703/runat_03062003_161407_22827/app_query.fasta.1.1870
-DB-SPTREMBL_21 -OPM=fastan -SUFFIX=rspt -MINMATCH=0.1 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptio -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09938703.ecgn.1.1.192.0runat_03062003_161407_22827 -NCPD=6 -ICPD=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-XGAPEXT=7 -XGAPOP=10 -XGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

SPTREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_nhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Query | Score | Match | Length | DB | ID | Description |
|------------|--------|-------|-------|--------|--------|----|--------------------|
| 1 | 1729.5 | 65.8 | 352 | 6 | Q95NC5 | | Q95nc5 hylobates s |

| | | | | | | |
|----|--------|------|-----|---|--------|--------------------|
| 2 | 1727.5 | 65.7 | 352 | 6 | 018771 | 018771 pan troglod |
| 3 | 1727.5 | 65.7 | 352 | 6 | 018772 | 018772 pan troglod |
| 4 | 1726.5 | 65.7 | 352 | 6 | Q9XS99 | Q9xs99 gorilla gor |
| 5 | 1725.5 | 65.7 | 352 | 6 | Q9TV50 | Q9tv50 pan troglod |
| 6 | 1722.5 | 65.5 | 352 | 6 | Q95NC0 | Q95nc0 hylobates m |
| 7 | 1721.5 | 65.5 | 352 | 6 | Q95NC7 | Q95nc7 nasalis lar |
| 8 | 1720.5 | 65.5 | 352 | 6 | Q95NC8 | Q95nc8 colobus pol |
| 9 | 1717.5 | 65.4 | 352 | 6 | Q97962 | Q97962 pygathrix a |
| 10 | 1716.5 | 65.3 | 352 | 6 | Q9XT14 | Q9xt14 colobus que |
| 11 | 1715.5 | 65.3 | 352 | 6 | Q95NC1 | Q95nc1 theropithec |
| 12 | 1715.5 | 65.3 | 352 | 6 | Q95NC6 | Q95nc6 trachypithe |
| 13 | 1715.5 | 65.3 | 352 | 6 | Q95NC3 | Q95nc3 miopithecus |
| 14 | 1713.5 | 65.2 | 352 | 6 | Q9XT13 | Q9xt13 papio anubi |
| 15 | 1712.5 | 65.2 | 352 | 6 | 018770 | 018770 pan troglod |
| 16 | 1712.5 | 65.2 | 352 | 6 | Q9TSK1 | Q9tsk1 cercopithec |
| 17 | 1712.5 | 65.2 | 352 | 6 | Q9TV49 | Q9tv49 cercocebus |
| 18 | 1711.5 | 65.1 | 352 | 6 | Q95ND2 | Q95nd2 mandrillus |
| 19 | 1711.5 | 65.1 | 352 | 6 | Q97975 | Q97975 macaca arct |
| 20 | 1710.5 | 65.1 | 352 | 6 | Q9XT12 | Q9xt12 cercopithec |
| 21 | 1709.5 | 65.0 | 352 | 6 | Q95ND1 | Q95nd1 mandrillus |
| 22 | 1707.5 | 65.0 | 352 | 6 | Q9TV42 | Q9tv42 cercopithec |
| 23 | 1707.5 | 65.0 | 352 | 6 | Q95NE8 | Q95ne8 ceropithec |
| 24 | 1706.5 | 64.9 | 352 | 6 | 077776 | 077776 cercocebus |
| 25 | 1706.5 | 64.9 | 352 | 6 | Q9MZA3 | Q9mza3 hylobates a |
| 26 | 1704.5 | 64.9 | 352 | 6 | Q95NE1 | Q95ne1 cercocebus |
| 27 | 1704.5 | 64.9 | 352 | 6 | Q95ND0 | Q95nd0 erythrocebu |
| 28 | 1703.5 | 64.8 | 352 | 6 | Q9TXQ0 | Q9txq0 cercopithec |
| 29 | 1701.5 | 64.7 | 352 | 6 | Q9TV93 | Q9tv93 macaca arct |
| 30 | 1700.5 | 64.7 | 352 | 6 | Q9TV46 | Q9tv46 cercopithec |
| 31 | 1698.5 | 64.6 | 352 | 6 | Q9BGN5 | Q9bgn5 ceropithec |
| 32 | 1698.5 | 64.6 | 352 | 6 | Q9XS35 | Q9xs35 macaca neme |
| 33 | 1696.5 | 64.6 | 352 | 6 | Q9XT76 | Q9xt76 ceropithec |
| 34 | 1695.5 | 64.5 | 352 | 6 | Q9TV43 | Q9tv43 ceropithec |
| 35 | 1694.5 | 64.5 | 352 | 6 | Q9TV48 | Q9tv48 ceropithec |
| 36 | 1693.5 | 64.4 | 352 | 6 | Q9TV47 | Q9tv47 ceropithec |
| 37 | 1693.5 | 64.4 | 352 | 6 | Q9TSQ7 | Q9tsq7 ceropithec |
| 38 | 1690.5 | 64.3 | 352 | 6 | Q9MZA2 | Q9mza2 ceropithec |
| 39 | 1683.5 | 64.1 | 339 | 4 | Q9UBT9 | Q9ubt9 homo sapien |
| 40 | 1683.5 | 64.1 | 352 | 6 | Q9TV44 | Q9tv44 ceropithec |
| 41 | 1680.5 | 63.9 | 339 | 4 | Q9UN27 | Q9un27 homo sapien |
| 42 | 1680.5 | 63.9 | 339 | 4 | Q9UBJ7 | Q9ubj7 homo sapien |
| 43 | 1679.5 | 63.9 | 339 | 4 | Q9UN23 | Q9un23 homo sapien |
| 44 | 1679.5 | 63.9 | 352 | 6 | Q9BGN6 | Q9bgn6 ceropithec |
| 45 | 1678.5 | 63.9 | 339 | 4 | Q9UN28 | Q9un28 homo sapien |

ALIGNMENTS

| RESULT 1 | | ALIGNMENTS | |
|----------|---|------------|---------|
| Q95NC5 | PRELIMINARY; | PRT; | 352 AA. |
| AC | Q95NC5; | | |
| DT | 01-DEC-2001 (TREMBLrel. 19, Created) | | |
| DT | 01-DEC-2001 (TREMBLrel. 19, Last sequence update) | | |
| DT | 01-MAR-2002 (TREMBLrel. 20, Last annotation update) | | |
| DE | C-C chemokine receptor 5. | | |
| GN | CCR5. | | |
| OS | Hylobates syndactylus (Siamang) (Sympalangus syndactylus). | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | |
| OC | Mammalia; Eutheria; Primates; Catarrhini; Hylobatidae; Hylobates. | | |
| OX | NCBI_TaxID=9590; | | |
| RN | [1] | | |
| RP | SEQUENCE FROM N.A. | | |
| RA | Zhang Y., Ryder O.A., Zhang Y.; | | |
| RT | "Sequence comparison of the CCR5 gene in primates and primate | | |
| RT | phylogeny." | | |
| RL | Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases. | | |
| DR | EMBL; AF177884; AAK43367.1; - | | |
| DR | InterPro: IPR000276; GPCR_Rhodpsn. | | |
| DR | Pfam: PF00001; 7tm1.1; 1. | | |
| DR | PROSITE: PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1. | | |
| DR | PROSITE: PS50262; G_PROTEIN_RECEP_F1_2; 1. | | |
| KW | Receptor. | | |

SQ SEQUENCE 352 AA; 40508 MW; F4F64B3AD5AF658A CRC64;

Alignment Scores:
 Pred. No.: 2,19e-162 Length: 352
 Score: 1729.50 Matches: 336
 Percent Similarity: 96.03% Conservative: 3
 Best Local Similarity: 95.18% Mismatches: 2
 Query Match: 65.81% Indels: 12
 DB: 6 Gaps: 1

US-09-938-703-3 (1-1442) x Q95NC5 (1-352)

QY 240 ATGGATTATCAAGTGTCAAGTCCATCTATGACATCAATATTATATACATCGAGCCCTGC 299
 Db 1 MetAspTyrGlnValSerProThrTyrAspIleAspTyrTyrThrSerGluProCys 20
 QY 300 CAAAAATCAATGTGAAGCAAAATCGAGCCGCTCTCCCTCCGCTCTACTACTGCTGTG 359
 Db 21 GlnLysIleAsnValLysGlnIleAlaAlaArgLeuProLeuTyrSerLeuVal 40
 QY 360 TTCATCTTTGGTTTGGGGCAACATGCTGCTCATCTCATCTGATCAATCACTGCAAAAGG 419
 Db 41 PheIlePheGlyPheValGlyAsnMetLeuValIleLeuIleLeuIleAsnCysLysArg 60
 QY 420 CTGAAGAGCATGACTGACATCTACCTGCTCAACCTGGCATCTCTGACCTGTTTTCCTT 479
 Db 61 LeuLysSerMetThrAspIleTyrLeuLeuAsnLeuAlaIleSerAspLeuPheLeu 80
 QY 480 CTTACTGTCCTTCCTGGCTCCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 539
 Db 81 LeuThrValProPheThrAlaHisTyrAlaAlaAlaGlnTyrAspPheGlyAsnThrMet 100
 QY 540 TGTCAACTCTTGACAGGCTCTATTATATAGGCTTCTCTGGAATCTTCTCATCATC 599
 Db 101 CysGlnLeuLeuThrGlyLeuTyrPheIleGlyPheSerGlyIlePhePheIleIle 120
 QY 600 CTCCTGACATTCGATAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 659
 Db 121 LeuLeuThrIleAspArgTyrLeuAlaIleValHisAlaValPheAlaLeuLysAlaArg 140
 QY 660 AGCTGCACCTTTGGGTGTGACAAGTGTGATCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 719
 Db 141 ThrValThrPheGlyValValThrSerValIleThrTyrValValAlaValPheAlaSer 160
 QY 720 CTCCTGACATTCCTTACAGATCTCAAAAGAGGTCTCTATTCATCTGACCTGCTGCT 779
 Db 161 LeuProGlyIleIlePheThrArgSerGlnLysGluGlyLeuHisTyrThrCysSerSer 180
 QY 780 CATTTTCCATAC-----ATTAAGATAGTCAT 806
 Db 181 HisPheProTyrSerGlnTyrGlnPheThrLysAsnPheGlnThrLeu-LysIleValIle 200
 QY 807 CTTGGGCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 866
 Db 200 eLeuGlyLeuValLeuProLeuValMetValIleCysTyrSerGlyIleLeuLysTh 220
 QY 867 TCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 926
 Db 220 rLeuLeuArgCysArgAsnGluLysLysArgHisArgAlaValArgLeuIlePheThrIle 240
 QY 927 CATGATTGTTTATTTCTCTCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 986
 Db 240 eMetIleValTyrPheLeuPheThrAlaProTyrAsnIleValLeuLeuLeuAsnThrPh 260
 QY 987 CCAGGAATCTTTGGCTCAATAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1046
 Db 260 eGlnGluPhePheGlyLeuAsnAsnCysSerSerSerAsnArgLeuAspGlnAlaMetG1 280
 QY 1047 GGTGACAGACTCTTGGGATGACGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1106
 Db 280 nValThrGluThrLeuGlyMetThrHisCysCysIleAsnProIleIleTyrAlaPheVa 300
 QY 1107 CGGGGAGAAGTTCAGAACTACCTCTTAGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1166

Db 300 lGlyGluLysPheArgAsnTyrLeuLeuValPhePheGlnLysHisIleAlaLysHisPh 320
 QY 1167 CTGCAAAATCCTGTTCTATTTTCCAGCAAGAGCTCCGAGCAGCAAGCTCACTTTTACAC 1226
 Db 320 ecLysLysCysCysSerIlePheGlnGlnAlaProGluArgAlaSerSerValTyrTh 340
 QY 1227 CCGATCCACTGGGAGCAGCAAAATATCTGTGGGCTTG 1263
 Db 340 rArgSerThrGlyGluGlnGluIleSerValGlyLeu 352

RESULT 2
 018771 PRELIMINARY; PRT; 352 AA.

AC 018771;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE CCR5 receptor (fragment).
 GN CCR5.
 OS Pan troglodytes (Chimpanzee).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
 OX NCBI_TaxID=9598;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CHC85-141a;
 RA Zhang L., Carruthers C.D., He T., Huang Y., Cao Y., Wang G., Hahn B.,
 RT "HIV-1 subtypes, co-receptor usage, and CCR5 polymorphism";
 RL AIDS Res. Hum. Retroviruses 0:0-0(1997).
 DR EMBL, AF011539; RAB65739.1;
 DR InterPro; IPR000276; GPCR_Rhodpsn.
 DR Pfam; PF00001; 7tm_1; 1.
 DR PRINTS; PR00237; GPCRHHODOPSN.
 DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
 DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
 KW Receptor.
 FT NON_TER. 352 352
 SQ SEQUENCE 352 AA; 40466 MW; 3FFFACT7ABAE1D4FB CRC64;

Alignment Scores:
 Pred. No.: 3,45e-162 Length: 352
 Score: 1727.50 Matches: 336
 Percent Similarity: 96.32% Conservative: 4
 Best Local Similarity: 95.18% Mismatches: 1
 Query Match: 65.73% Indels: 12
 DB: 6 Gaps: 1

US-09-938-703-3 (1-1442) x 018771 (1-352)

QY 240 ATGGATTATCAAGTGTCAAGTCCATCTATGACATCAATATTATATACATCGAGCCCTGC 299
 Db 1 MetAspTyrGlnValSerProThrTyrAspIleAspTyrTyrThrSerGluProCys 20
 QY 300 CAAAAATCAATGTGAAGCAAAATCGAGCCGCTCTCCCTCCGCTCTACTACTGCTGTG 359
 Db 21 GlnLysIleAsnValLysGlnIleAlaAlaArgLeuProLeuTyrSerLeuVal 40
 QY 360 TTCATCTTTGGTTTGGGGCAACATGCTGCTCATCTCATCTGATCAATCACTGCAAAAGG 419
 Db 41 PheIlePheGlyPheValGlyAsnMetLeuValIleLeuIleLeuIleAsnCysLysArg 60
 QY 420 CTGAAGAGCATGACTGACATCTACCTGCTCAACCTGGCATCTCTGACCTGTTTTCCTT 479
 Db 61 LeuLysSerMetThrAspIleTyrLeuLeuAsnLeuAlaIleSerAsnLeuPheLeu 80
 QY 480 CTTACTGTCCTTCCTGGCTCCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 539
 Db 81 LeuThrValProPheThrAlaHisTyrAlaAlaAlaGlnTyrAspPheGlyAsnThrMet 100
 QY 540 TGTCAACTCTTGACAGGCTCTATTATATAGGCTTCTCTGGAATCTTCTCATCATC 599


```

QY 1047 GGTGACAGAGCTTGGGATGACCGCCTGCTGCATCAACCCCATCATCTATGCCCTTGT 1106
Db 280 nValThrGluThrLeuGlyMetThrHisCysCysIleAsnProIleIleTyrAlaPheVa 300
QY 1107 CGGGAGAGAGTTCAGAAATACCTTCTAGTCTTCTTCCAAAAGACACATTGCCAAAGCCTT 1166
Db 300 IGlyGluLysPheArgAsnTyrLeuLeuValPhePheGlnLysHisIleAlaLysArgPh 320
QY 1167 CTGCAATGCTGTCTATTTTCCAGCAAGAGCTCCGAGGAGCAAGCTCAGTTTACAC 1226
Db 320 eCysLysCysSerIlePheGlnGlnGluAlaProGluArgAlaSerSerValTyrTh 340
QY 1227 CCGATCCACTGGGAGAGCAAAATATCTGTGGGCTTG 1263
Db 340 rArgSerThrGlyGluGlnGluIleSerValGlyLeu 352

RESULT 4
Q9XS99
ID Q9XS99 PRELIMINARY; PRT; 352 AA.
AC Q9XS99;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE CC chemokine receptor 5.
GN CCR5.
OS Gorilla gorilla (gorilla).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
OX NCBI_TaxID=9593;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GORILLACCR;
RX MEDLINE=99210133; PubMed=10195758;
RA Saksena N.K., Wang B., Novembre F.J., Bolton W., Smit T., Lal R.B.;
RT "Species-specific changes in the CCR5 gene from African and Asian
nonhuman primates.";
RL AIDS Res. Hum. Retroviruses 15:479-483(1999).
DR EMBL; AF105291; AAD20560.1; -.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PR00237; GPCRHOOPS.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 352 AA; 40529 MW; 1B568C68FE2E7AD0 CRC64;

Alignment Scores:
Pred. No.: 4,34e-162 Length: 352
Score: 1726.50 Matches: 335
Percent Similarity: 96.03% Conservative: 4
Best Local Similarity: 94.90% Mismatches: 2
Query Match: 65.70% Indels: 12
DB: 6 Gaps: 1

US-09-938-703-3 (1-1442) x Q9XS99 (1-352)
QY 240 ATGGATTATCAAGTGTCAAGTCCCAATCATGACATCAATATTATACATCGGAGCCCTG 299
Db 1 MetAspTyrGlnValSerSerProThrTyrAspIleAspTyrThrSerGluProCys 20
QY 300 CAANAATCAATGTGAAGCAATCGACGCCCTCGCTCCGCTCTACTACTGCTGTG 359
Db 21 GlnLysThrAsnValLysGlnIleAlaAlaArgLeuLeuProProLeuTyrSerLeuVal 40
QY 360 TTCATCTTTGTTTGGGCAACATGCTGGTCATCTCCATCCCTGATGAATGCAAGG 419
Db 41 PheIlePheGlyPheValGlyAsnMetLeuValIleLeuIleLeuLeuAsnCysLysArg 60
QY 420 CTGAAGAGCATGACTGACATCTACCTGCTCAACTGGCCATCTCTGACCTGTTTTCCT 479
Db 61 LeuLysSerMetThrAspIleTyrLeuLeuAsnLeuAlaIleSerAspLeuPheLeu 80
QY 480 CTTACTGTCCCTTCTGGGCTCACTATGCTGCGGCCAGTGGGACTTTTGAATACAATG 539

```

```

Db 81 LeuThrValProPheThrAlaHisTyrAlaAlaGlnTyrAspPheGlyAsnThrMet 100
QY 540 TGTCAACTCTTACAGGCTCTATTTATAGGCTTCTTCTGGAACTCTTCTTCATCATC 599
Db 101 CysGlnLeuLeuThrGlyLeuTyrPheIleGlyPheSerGlyIlePhePheIleIle 120
QY 600 CTCCTGCAATCGATPAGTACCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 659
Db 121 LeuLeuThrIleAspArgTyrLeuAlaIleValHisAlaValPheAlaLeuLysAlaArg 140
QY 660 ACSTGTCACCTTTGGGTGGTGCACAACTGATCATCTGGGTGGTGGTGGTGGTGGTGGT 719
Db 141 ThrValThrPheGlyLeuValThrSerValIleThrTrpValValAlaValPheAlaSer 160
QY 720 CTCGAGGAATCATCTTTACCATCTCAAAAAGAGTCTTCAATACACCTGAGCTCT 779
Db 161 LeuProGlyIleIlePheThrArgSerGlnLysGluGlyLeuHisTyrThrCysSerSer 180
QY 780 CATTTCCTATAC-----ATTAAAGATAGTTCAT 806
Db 181 HisPheProTyrSerGlnTyrGlnPheTrpLysAsnPheGlnThrLeu-LysIleValI 200
QY 807 CTTGGGCTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 866
Db 200 eLeuGlyLeuValLeuProLeuLeuValMetValIleCysTyrSerGlyIleLeuLysTh 220
QY 867 TCTGCTTCGCTGCTGAATGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 926
Db 220 rLeuLeuArgCysArgAsnGlnLysLysArgHisArgAlaValArgLeuIlePheThrI 240
QY 927 CATCATGTTTATTTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 986
Db 240 eMetIleValTyrPheLeuPheThrAlaProTyrAsnIleValLeuLeuLeuAsnThrPh 260
QY 987 CCAGGAATCTTTGGCTGCTGAATATGTCAGTAGTCTTAACAGGTTGGACCAAGCTATGCA 1046
Db 260 eGlnGluPhePheGlyLeuAsnAsnCysSerSerSerAsnArgLeuAspGlnAlaMetGl 280
QY 1047 GGTGACAGAGCTCTTGGGATGACGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1106
Db 280 nValThrGluThrLeuGlyMetThrHisCysCysIleAsnProIleIleTyrAlaPheVa 300
QY 1107 CGGGAGAGAGTTCAGAAATACCTTCTAGTCTTCTTCCAAAAGACACATTGCCAAAGCCTT 1166
Db 300 IGlyGluLysPheArgAsnTyrLeuLeuValPhePheGlnLysHisIleAlaLysArgPh 320
QY 1167 CTGCAATGCTGTCTATTTTCCAGCAAGAGCTCCGAGGAGCAAGCTCAGTTTACAC 1226
Db 320 eCysLysCysCysSerIlePheGlnGlnGluAlaProGluArgAlaSerSerValTyrTh 340
QY 1227 CCGATCCACTGGGAGAGCAAAATATCTGTGGGCTTG 1263
Db 340 rArgSerThrGlyGluGlnGluIleSerValGlyLeu 352

RESULT 5
Q9TW50
ID Q9TW50 PRELIMINARY; PRT; 352 AA.
AC Q9TW50;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE CC chemokine receptor type 5.
GN CCR5.
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX NCBI_TaxID=9598;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1410;
RX MEDLINE=99335215; PubMed=10408730;
RA Muller-Trutwin M.C., Corbet S., Hansen J., Georges-Courbot M.C.,

```


RA Diop O., Rigoulet J., Barre-Sinoussi F., Fomsgaard A.:
 RT "Mutations in CCR5-coding sequences are not associated with HIV
 carrier status in African nonhuman primates.";
 RT AIDS Res. Hum. Retroviruses 15:931-939(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-1410;
 RA Mueller-Trutwin M.C., Corbet S., Hansen J., Diop O.,
 RA Georges-Courbot M.-C., Barre-Sinoussi F., Fomsgaard A.;
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF035214; AAD44007.1;
 DR InterPro: IPR000276; GPCR_Rhodpsn.
 DR Pfam: PF00001; 7tm1.1;
 DR PRINTS; PR00237; GPCRHOOPS.
 DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
 DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
 KW Receptor.
 SQ SEQUENCE 352 AA; 40481 MW; 2578A0F2C07AA65 CRC64;
 Alignment Scores:
 Pred. No.: 5,45e-162 Length: 352
 Score: 1725.50 Matches: 334
 Percent Similarity: 96.32% Conservative: 6
 Best Local Similarity: 94.62% Mismatches: 1
 Query Match: 65.66% Indels: 12
 DB: 6 Gaps: 1

US-09-938-703-3 (1-1442) x Q9NV50 (1-352)

QY 240 ATGGATTATCAAGTCTCAAGTCCAACTATGATGATCAATATTATATACATCGGAGCCCTGC 299
 Db 1 MetAspTyrGlnValSerProThrTyrAspIleAspTyrThrSerGluProCys 20
 QY 300 CAAAATCAATGTGAAGCAATGCGAGCCGCTCGCTCGCTCTACTACTGCTG 359
 Db 21 GlnTyrLeuValSerGlnValGlnValGlnValGlnValGlnValGlnVal 40
 QY 360 TTCACTCTTGTGTTGCGCAACATGCTGCTCATCTCTCTGATCACTGCAAAAGG 419
 Db 41 PheIlePheGlyPheValGlyAsnIleValValIleValIleValIleValIleVal 60
 QY 420 CTGAAGACATGATGACATCTACCTGCTCAACCTGGCCATCTCTGACCTGTTTCTCT 479
 Db 61 LeuTyrSerMetThrAspIleTyrLeuLeuAsnLeuAlaIleSerAspLeuPheLeu 80
 QY 480 CTACTGTCCTCTCGGCTCATCTGCTGCGCCGACCTGGGACTTGGAAATCAATG 539
 Db 81 LeuThrValProPheThrAlaHisTyrAlaAlaAlaGlnTrpAspPheGlyAsnThrMet 100
 QY 540 TGCTCACTCTGTGACGGCTCTATTTATAGGCTTCTCTGGAATCTTCTTCATCATC 599
 Db 101 CysGlnLeuLeuThrGlyLeuTyrPheIleGlyPhePheSerGlyIlePhePheIle 120
 QY 600 CTCCTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 659
 Db 121 LeuLeuSerIleAspArgTyrLeuAlaIleValHisAlaValPheAlaLeuLysAlaArg 140
 QY 660 ACGGTCACCTTGGGCTGGTGAAGTGTGATCACTTGGGCTGGGCTGGGCTGGGCT 719
 Db 141 ThrValThrPheGlyValValThrSerValIleThrTrpValValAlaValPheAlaSer 160
 QY 720 CTCCTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 779
 Db 161 LeuProGlyIleIlePheThrArgSerGlnLysGlnGlyLeuHisTyrThrCysSerSer 180
 QY 780 CATTTTCCATC-----ATTAAGATAGTCAT 806
 Db 181 HisPheProTyrSerGlnTyrGlnPheThrLysAsnPheGlnThrLeu-LysIleValI 200
 QY 807 CTGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 866
 Db 200 eLeuGlyLeuValLeuProLeuValMetValIleCysTyrSerGlyIleLeuLysTh 220

QY 867 TCTGCTTCGGTCTCGAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 926
 Db 220 rLeuLeuArgCysArgAsnGlnLysLysArgHisArgAlaValArgLeuIlePheThrI 240
 QY 927 CATGATTGTTATTTCTCTCTCTGGGCTCCCTCAACATGCTCTCTCTCTGACACCTT 986
 Db 240 eMetIleValIyrPheLeuPheThrPheThrPheThrPheThrPheThrPheThrP 260
 QY 987 CCAGGAATTTCTTGGCCTGAATTAATTCAGTAGCTCTAAGAGTTGGACCAAGCTATGCA 1046
 Db 260 eGlnGluPhePheGlyLeuAsnAsnCysSerSerSerAsnArgLeuAspGlnAlaMetG 280
 QY 1047 GTGACAGAGACTTGTGGATGACGACTGTGATCATCAACCCCATCATCATGCTTTGT 1106
 Db 280 nValThrGluThrLeuGlyMetThrHisCysCysIleAsnProIleIleTyrAlaPheVa 300
 QY 1107 CGGGAGAAAGTTCAAGAACTACCTCTAGTCTCTCTCCAAAAGCAGCATGCGCAAGCTT 1166
 Db 300 LGlyLysPheArgAsnTyrLeuLeuValPhePheGlnLysHisIleAlaLysArgPh 320
 QY 1167 CTGCAAAATGCTGTTCTATTATTTCCAGCAAGGCTCCCGAGCGAGCAAGCTCAGTTTAC 1226
 Db 320 eCysLysCysCysSerIlePheGlnGlnGlnAlaProGluArgAlaSerSerValThr 340
 QY 1227 CCGATCCACTGGGAGCAGGAGAAATATCTGTGGGCTTG 1263
 Db 340 rArgSerThrGlyGluGlnGluIleSerValGlyLeu 352

RESULT 6

Q95NC0
 ID Q95NC0 PRELIMINARY; PRT; 352 AA.
 AC Q95NC0;
 DT 01-DEC-2001 (TEMBLrel. 19, Created)
 DT 01-DEC-2001 (TEMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TEMBLrel. 20, Last annotation update)
 DE C-C chemokine receptor 5.
 GN CCR5.
 OS Hylobates moloch (silvery gibbon).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hylobatidae; Hylobates.
 OX NCBI_TaxID=81572;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Zhang Y., Ryder O.A., Zhang Y.;
 RT "Sequence comparison of the CCR5 gene in primates and primate
 RT Phylogeny.";
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF177899; AAK43382.1;
 DR InterPro: IPR000276; GPCR_Rhodpsn.
 DR Pfam: PF00001; 7tm1.1;
 DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
 DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
 KW Receptor.
 SQ SEQUENCE 352 AA; 40436 MW; 9623CA98340CF274 CRC64;

Alignment Scores:

Pred. No.: 1,08e-161 Length: 352
 Score: 1722.50 Matches: 335
 Percent Similarity: 95.75% Conservative: 3
 Best Local Similarity: 94.90% Mismatches: 3
 Query Match: 65.54% Indels: 12
 DB: 6 Gaps: 1

US-09-938-703-3 (1-1442) x Q95NC0 (1-352)

QY 240 ATGGATTATCAAGTCTCAAGTCCAACTATGATGATCAATATTATATACATCGGAGCCCTGC 299
 Db 1 MetAspTyrGlnValSerProThrTyrAspIleAspTyrThrSerGlyProCys 20
 QY 300 CAAAATCAATGTGAAGCAAAATCGCAGCCGCTCGCTCGCTCGCTCGCTCGCTCGCTG 359
 Db 21 GlnLysLeuValLysGlnIleAlaAlaArgLeuLeuProLeuTyrSerLeuVal 40

| | | | |
|----|------|--|------|
| db | 200 | eLeuGlyLeuValLeuProLeuValMetValIleCysTyrSerGlyLeuLysTh | 220 |
| QY | 867 | TCTGCTTCGGTGTGAAATGAGAAAGACGAGCAGGGCTGTGAGGCTTATCTTCAACAT | 926 |
| Db | 220 | rLeuLeuArgCysArgAsnGluLysLysArgHisArgAlaValArgLeuIlePheThrI | 240 |
| QY | 927 | CATGATTGTTTATTTCTCTTCCTGGGCTCCCTACACATTGCTCTTCTCTGACACCTT | 986 |
| Db | 240 | eMetIleValTyrPheLeuPheTrpAlaProTyrAsnIleValLeuLeuLeuAsnTrpPh | 260 |
| QY | 987 | CCAGGAATCTTTGGCGCTGAATTAATTCAGTAGCTCTAACAGTTGGACCAAGCTATGCA | 1046 |
| Db | 260 | eGluGluPhePheGlyLeuAsnAsnCysSerSerAsnArgLeuAspGlnAlaMeG | 280 |
| QY | 1047 | GGTGACAGAGACTCTGGGATGAGCGACTGCTGCATCAACCCCATCATCTATGCGCTTGT | 1106 |
| Db | 280 | nValThrGluThrLeuGlyMetThrHisCysCysIleAsnProIleIleTyrAlaPheVa | 300 |
| QY | 1107 | CGGGAGAGAGTTCAGAACTACTCTTAGTCTTCTTCCAAAGACACATTGCCAAAGCCTT | 1166 |
| Db | 300 | IglYgluLysPheArgAsnTyrLeuLeuValPhePheGlnIleYsHisIleAlaLysArgph | 320 |
| QY | 1167 | CTGCAATGCTGTTCTATTTTCAGCAAGAGGCTCCGAGCGAGCAAGCTCAGTTTACAC | 1226 |
| Db | 320 | eCysLysCysCysSerIlePheGlnGlnAlaProGluArgAlaSerSerValTyrTh | 340 |
| QY | 1227 | CGATCCACTGGGGAGCAGCAAAATATCTGTGGGCTTG | 1263 |
| Db | 340 | rArgSerThrGlyGluGlnGluThrSerValGlyLeu | 352 |

RESULT 10

Q9XTI4

| | | | | |
|-----|--|---|-----------|------------------------|
| ID | Q9XTI4 | PRELIMINARY; | PRT; | 352 AA. |
| IC | Q9XTI4: | | | |
| DT | 01-NOV-1999 | (TrEMBLrel. 12, Created) | | |
| DDT | 01-NOV-1999 | (TrEMBLrel. 12, Last sequence update) | | |
| DDT | 01-DEC-2001 | (TrEMBLrel. 19, Last annotation update) | | |
| DE | CC | chemokine receptor 5. | | |
| GN | CCR5 | | | |
| OS | Colobus quereza (Black-and-white colobus monkey). | | | |
| OC | Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi; | | | |
| OC | Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae; Colobinae; | | | |
| OC | Colobus. | | | |
| OX | NCBI_TaxID=33548; | | | |
| XP | SEQUENCE FROM N.A. | | | |
| RP | Spearman P.W., Mouru D.N., Graham B.S.; | | | |
| RT | *Differential Utilization of CCR5 Molecules from Three East African | | | |
| RT | Simian Species by the HIV-1 Envelope Glycoprotein.; | | | |
| RL | Submitted (APR-1999) to the EMBL/Genbank/DBJ databases. | | | |
| RL | EMBL; AF141639; AAD32684.1; - | | | |
| DR | InterPro: IPR000276; GPCR_Rhodpsn. | | | |
| DR | Pfam: PF00001; 7tm_1; 1. | | | |
| DR | PRINTS: PR00237; GPCRHHODOPSN. | | | |
| DR | PROSITE: PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1. | | | |
| DR | PROSITE: PS50262; G_PROTEIN_RECEP_F1_2; 1. | | | |
| DR | Receptor. | | | |
| SK | SEQUENCE | 352 AA; | 40550 MW; | 9B078EF04D3DB36 CRC64; |

Alignment Scores:

| | | | |
|------------------------|-----------|---------------|-----|
| Pred. No.: | 4,24e-161 | Length: | 352 |
| Score: | 1716.50 | Matches: | 331 |
| Percent Similarity: | 95.75% | Conservative: | 7 |
| Best Local Similarity: | 93.77% | Mismatches: | 3 |
| Query Match: | 65.32% | Indels: | 12 |
| DB: | 6 | Gaps: | 1 |

US-09-938-703-3 (1-1442) x Q9XTI4 (1-352)

| | | | |
|----|-----|--|-----|
| QY | 240 | ATGGATTATCAAGTGTCAAGTCCAACTATGACATCAATATTATATACATCGGACCGCTGC | 299 |
| Db | 1 | MetAspTyrGlnValSerSerProThrTyrAspIleAsnTyrThrSerGluProCys | 20 |

QY 300 CAAAAATCAATGTCAGCAAAATCGAGCCCGCTCTGCTCCCTGCTCTACTACTGGTG 359
 Db 21 GlnlyValAsnValysGlnlleAlaAlaArgLeuLeuProLeuTyrSerLeuVal 40
 QY 360 TTCATCTTTGGTTTGGGCAACATGCTGGTCACTCTCATCTGATAAATGCAAAAGG 419
 Db 41 PheillePheGlyPheValGlyAsnilleLeuValValleulleLeulleAsnCyslySarg 60
 QY 420 CTGAAGAGCATGACTGACATCTACTCTCAACCTGGCCATCTCTGACCTGTTTCCCT 479
 Db 61 LeuLysSerMetThrAspilleTyrLeuLeuAsnLeuAlaIleSerAspLeuPheLeu 80
 QY 480 CTTACTGTCCTTCTGGGCTCACTATGCTCCGCCAGTGGGACTTTGGAATACATG 539
 Db 81 LeuThrValProPheThrAlaHisTyrAlaAlaArgLeuLeuValPheAlaSer 100
 QY 540 TGTCAACTCTTGACAGGGCTTATTTATAGGCTTCTCTGGAATCTTCTTCATCATC 599
 Db 101 CysGlnLeuLeuThrGlyLeuTyrPheilleGlyPhePheSerGlyLePhePheille 120
 QY 600 CTCCTGACAAATGATAGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 659
 Db 121 LeuLeuThrIleAspArgTyrLeuAlaIleValHisAlaValPheAlaLeuLysAlaArg 140
 QY 660 ACGGTCACTTTGGGCTGGTGCACAGTGTGATCACTTGGTGGTGGTGGTGGTGGTCT 719
 Db 141 ThrAlaThrPheGlyValValThrSerValIleThrValValAlaValPheAlaSer 160
 QY 720 CTCCTGAGCAATGATAGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 779
 Db 161 LeuProGlyIleIlePheThrArgSerGlnArgGlyLeuHisTyrThrCysSerSer 180
 QY 780 CATTTTCCATAC-----ATTAAGATAGTATCAT 806
 Db 181 HisPheProTyrSerGlnTyrGlnPheThrLysAsnPheGlnThrLeu-LysIleValI 200
 QY 807 CTTGGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 866
 Db 200 eueGlyLeuValLeuProLeuLeuValMetValIleCysTyrSerGlyIleLeuLysTh 220
 QY 867 TCTGCTTCGGTTCGAAATGAGAGAGAGGACAGGCTGTGAGGCTTATCTTCAACAT 926
 Db 220 rleuLeuArgCysArgAsnGlnLysLysArgHisArgAlaValArgLeuIlePheThrI 240
 QY 927 CATGATGTTTATTTCTTCTTCTGGCTCCCTGACAGATGCTCTCTCTCTCTGACACCT 986
 Db 240 emetIleValTyrPheLeuPheThrAlaProTyrAsnIleValLeuLeuAsnThrPh 260
 QY 987 CCAGGAATCTTTGGCCGTAATAATTCAGTAGCTTAACAGGTTGGACCAAGCTATGCA 1046
 Db 260 eGlnGlnPhePheGlyLeuAsnAsnCysSerSerSerSerSerSerSerSerSerSer 1280
 QY 1047 GGTGACAGACACTCTTGGGATGAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1106
 Db 280 nValThrGluThrLeuGlyMetThrHisCysCysIleAsnProIleIleTyrAlaPheVa 300
 QY 1107 CGGGGAGAAGTTCAGAACTACTCTTACTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1166
 Db 300 lGlyGlyLysPheArgAsnTyrLeuLeuValPhePheGlnIlySHisIleAlaLysArgPh 320
 QY 1167 CTGCAAAATGCTGTTCTTATTTTCCAGCAAGAGCTCCCGAGGAGCAAGCTGATTTAC 1226
 Db 320 ecyslyCysCysArgIlePheGlnGlnGlnAlaProGluArgAlaSerSerValTyrTh 340
 QY 1227 CGATPCCACTGGGGAGCAGGAATATCTGTGGGCTG 1263
 Db 340 rArgSerThrGlyGluGlnGluIleSerValGlyLeu 352
 RESULT 11
 Q95NCL
 ID Q95NCL PRELIMINARY; PRT; 352 AA.
 AC Q95NCL;

DT 01-DEC-2001 (TRENBLrel. 19, Created)
 DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
 DE C-C chemokine receptor 5.
 GN CCR5.
 OS Theropithecus gelada (Gelada baboon).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
 OC Cercopitheinae; Theropithecus.
 RX NCBI_TaxID=9565;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Zhang Y., Ryder O.A., Zhang Y.;
 RT "Sequence comparison of the CCR5 gene in primates and primate phylogeny."
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AFI77891; AAK43374.1;
 DR InterPro; IPR000276; GPCR_Rhodpsn.
 DR Pfam; PF00001; 7tm.1; 1
 DR PROSITE; PS00237; G-PROTEIN_RECEP_FL_1; UNKNOWN_1.
 DR PROSITE; PS00262; G-PROTEIN_RECEP_FL_2; 1.
 KW Receptor.
 SQ SEQUENCE 352 AA; 40475 MW; 4D3643B6C13E90D5 CRC64;
 Alignment Scores:
 Pred. No.: 4.24e-161 Length: 352
 Score: 1716.50 Matches: 332
 Percent Similarity: 95.75% Conservative: 6
 Best Local Similarity: 94.05% Mismatches: 3
 Query Match: 65.32% Indels: 12
 DB: 6 Gaps: 1
 US-09-938-703-3 (1-1442) x Q95NCL (1-352)
 QY 240 ATGGATTATCAAGTCAAGTCAATCTATGATCATCAATATTATATCATCGAGCCCTGC 299
 Db 1 MetAspTyrGlnValSerSerProThrTyrAspIleAspTyrTyrThrSerGluProCys 20
 QY 300 CAAAAATCAATGTCAGCAAAATCGAGCCCGCTCTGCTCCCTGCTCTACTACTGGTG 359
 Db 21 GlnLysIleAsnValLysGlnlleAlaGlyArgLeuLeuProLeuTyrSerLeuVal 40
 QY 360 TTCATCTTTGGTTTGGGCAACATGCTGGTCACTCTCATCTGATAAATGCAAAAGG 419
 Db 41 PheillePheGlyPheValGlyAsnilleLeuValValleulleLeulleAsnCyslySarg 60
 QY 420 CTGAAGAGCATGACTGACATCTACTCTCAACCTGGCCATCTCTGACCTGTTTCCCT 479
 Db 61 LeuLysSerMetThrAspilleTyrLeuLeuAsnLeuAlaIleSerAspLeuPheLeu 80
 QY 480 CTTACTGTCCTTCTGGGCTCACTATGCTCCGCCAGTGGGACTTTGGAATACATG 539
 Db 81 LeuThrValProPheThrAlaHisTyrAlaAlaArgLeuLeuValPheAlaSer 100
 QY 540 TGTCAACTCTTGACAGGGCTTATTTATAGGCTTCTCTGGAATCTTCTTCATCATC 599
 Db 101 CysGlnLeuLeuThrGlyLeuTyrPheilleGlyPhePheSerGlyLePhePheille 120
 QY 600 CTCCTGACAAATGATAGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 659
 Db 121 LeuLeuThrIleAspArgTyrLeuAlaIleValHisAlaValPheAlaLeuLysAlaArg 140
 QY 660 ACGGTCACTTTGGGCTGGTGCACAGTGTGATCACTTGGTGGTGGTGGTGGTGGTCT 719
 Db 141 ThrValThrPheGlyValValThrSerValIleThrValValAlaValPheAlaSer 160
 QY 720 CTCCAGGAATCATCTTTTACAGATCTCAAAAGAGGCTTCTTCACTACCTGACCTCT 779
 Db 161 LeuProGlyIleIlePheThrArgSerGlnArgGlyLeuHisTyrThrCysSerSer 180
 QY 780 CATTTTCCATAC-----ATTAAGATAGTATCAT 806
 Db 181 HisPheProTyrSerGlnTyrGlnPheThrLysAsnPheGlnThrLeu-LysIleValI 200

RESULT 14

| | | | |
|---|---|---------------|---------|
| Q9XT13 | PRELIMINARY; | PRT; | 352 AA. |
| ID | Q9XT13 | | |
| AC | Q9XT13 | | |
| DT | 01-NOV-1999 (TREMBLrel. 12, Created) | | |
| DT | 01-NOV-1999 (TREMBLrel. 12, Last sequence update) | | |
| DT | 01-JUN-2001 (TREMBLrel. 17, Last annotation update) | | |
| DE | CC chemokine receptor 5. | | |
| DE | CC | | |
| GN | CCR5. | | |
| OS | Papio anubis (Olive baboon). | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | |
| OC | Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae; | | |
| OC | Cercopithecinae; Papio. | | |
| OX | NCBI_TaxID=9555; | | |
| RP | [1] | | |
| RN | SEQUENCE FROM N.A. | | |
| RA | Spearmant P.W., Mburu D.N., Graham B.S.; | | |
| RT | "Differential Utilization of CCR5 Molecules from Three East African | | |
| RT | Simian Species by the HIV-1 Envelope Glycoprotein."; | | |
| RL | Submitted (APR-1999) to the EMBL/GenBank/DBAJ databases. | | |
| DR | EMBL; AF141640; AAC32685.1; - | | |
| DR | InterPro; IPR000276; GPCR_Rhodpsn. | | |
| DR | Pfam; PF00001; 7tm1.1; | | |
| DR | PRINTS; PR00237; GPCRHHODPSN. | | |
| DR | PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1. | | |
| DR | PROSITE; PS02462; G_PROTEIN_RECEP_F1_2; 1. | | |
| KW | Receptor. | | |
| SQ | SEQUENCE 352 AA; 40489 MW; 0B47E337C1E2E1E CRC64; | | |
| Alignment Scores: | | | |
| Pred. No.: | 8.4e-161 | Length: | 352 |
| Score: | 1713.50 | Matches: | 332 |
| Percent Similarity: | 95.75% | Conservative: | 6 |
| Best Local Similarity: | 94.05% | Mismatches: | 3 |
| Query Match: | 65.20% | Indels: | 12 |
| DB: | 6 | Gaps: | 1 |
| US-09-938-703-3 (1-1442) x Q9XT13 (1-352) | | | |


```
QY 240 ATGGATATCAAGTGTCAAGTCCCAATCTATGACATCAATATATATATACATCGAGCCCTGC 299
Db 1 MetAspTyrGlnValSerSerProThrTyrAspIleAspTyrTyrThrSerGluProCys 20
QY 300 CAAAAATCAATGTGAAGCAATCGAGCCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 359
Db 21 GlnLysIleAsnValLysGlnIleAlaArgLeuLeuProLeuTyrSerLeuVal 40
QY 360 TTCTATCTTTGGTTTGGGCAACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 419
Db 41 PheIlePheGlyPheValGlnIleValIleValIleValIleValIleValIleValIle 60
QY 420 CTGAGAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 479
Db 61 LeuLysSerMetThrAspAsnTyrLeuLeuAsnLeuAlaIleSerAspLeuLeuPheLeu 80
QY 480 CTCTACTGCTCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 539
Db 81 LeuThrValProPheThrPalaHisTyrAlaAlaGlnTyrPaspPheGlyAsnThrMet 100
QY 540 TGTCAACTCTTGACAGGCTCTATTTTATAGGCTTCTCTGGAATCTTCTTCATCATC 599
Db 101 CysGlnLeuLeuThrGlyLeuTyrPheIleGlyPheSerGlyIlePhePheIleIle 120
QY 600 CTCTGTCACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 659
Db 121 LeuLeuThrIleAspArgTyrLeuAlaIleValHisAlaValPheAlaLeuLysAlaArg 140
QY 660 ACGGTCACTTTGGGGTGGTGAAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 719
Db 141 ThrValThrPheGlyValValThrSerValIleThrTrpValValAlaValPheAlaSer 160
QY 720 CTCCAGGATCATCTTTACAGATGATCAAAAAGAGTCTTATACCTGCGAGCTCT 779
Db 161 LeuProGlyIleIlePheThrArgSerGlnArgGluGlyLeuHisTyrThrCysSerSer 180
QY 780 CATTTTCCATAC-----ATTAAGATAGTATCAT 806
Db 181 HisPheProTyrSerGlnTyrGlnPheThrPlysAsnPheGlnThrLeu-LysIleValI 200
QY 807 CTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 866
Db 200 eLeuGlyLeuValLeuProLeuLeuValMetValIleCysTyrSerGlyIleLeuLysTh 220
QY 867 TCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 926
Db 220 rLeuLeuArgCysArgAsnGlnLysLysArgHisArgAlaValArgLeuIlePheThrIl 240
QY 927 CATGATTTGTTTATTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 986
Db 240 eMetIleValTyrPheLeuPheThrPalaProTyrAsnIleValLeuLeuLeuAsnThrPh 260
QY 987 CCAGGAATCTTTGGCTGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1046
Db 260 eGlnGluPhePheGlyLeuAsnAsnCysSerSerSerAsnArgLeuAspGlnAlaMetGl 280
QY 1047 GGTGACAGAGACTTTGGGATGAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1106
Db 280 nValThrGluThrLeuGlyMetThrHisCysLysIleAsnProIleIleThrAlaPheVa 300
QY 1107 CGGGAGAGTTCAGAAATACCTCTTAGTCTTCTTCCAAAAGCACATGCCCACACGCTT 1166
Db 300 lGlyGluLysPheArgAsnTyrLeuLeuValPhePheGlnLysHisIleAlaLysArgPh 320
QY 1167 CTGCAATGCTGTTCTATTTTCCAGCAGAGGCTCCGAGGAGCAGAGCTGATGATGATGATGAT 1226
Db 320 eCysLysCysSerIlePheGlnGlnGluAlaProGluArgAlaSerSerValTyrThr 340
QY 1227 CCGATCCACTGGGAGAGGAGAAATATCTGTGGGCTTG 1263
Db 340 rArgSerThrGlyGlnGlnGluIleSerValGlyLeu 352
```

RESULT 15

```
018770
ID 018770 PRELIMINARY; PRT; 352 AA.
AC 018770;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE CCR5 receptor (Fragment).
GN CCR5.
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX NCBI_TaxID=9598;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MACCR5-140A;
RA Zhang L., Carruthers C.D., He T., Huang Y., Cao Y., Wang G., Hahn B.,
RA "HIV-1 subtypes, co-receptor usage, and CCR5 polymorphism.";
RL AIDS Res. Hum. Retroviruses 0:0-0(1997).
DR EMBL; AF011538; AAB65738.1;
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1;
DR PRINTS; PR00237; GPCRHHODPSN.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE; PS00262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT NON_TER 352
SQ SEQUENCE 352 AA; 40523 MW; 4513DB983A28ACB2 CRC64;

Alignment Scores:
Pred. No.: 1,06e-160 Length: 352
Score: 1712.50 Matches: 331
Percent Similarity: 95.75% Conservative: 7
Best Local Similarity: 93.77% Mismatches: 3
Query Match: 65.16% Indels: 12
DB: Gaps: 1

US-09-938-703-3 (1-1442) x 018770 (1-352)
QY 240 ATGGATATCAAGTGTCAAGTCCCAATCTATGACATCAATATATATATACATCGAGCCCTGC 299
Db 1 MetAspTyrGlnValSerSerProThrTyrAspIleAspTyrTyrThrSerGluProCys 20
QY 300 CAAAAATCAATGTGAAGCAATCGAGCCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 359
Db 21 GlnLysIleAsnValLysGlnIleAlaArgLeuLeuProLeuTyrSerLeuVal 40
QY 360 TTCTATCTTTGGTTTGGGCAACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 419
Db 41 PheIlePheGlyPheValGlnIleValIleValIleValIleValIleValIleValIle 60
QY 420 CTGAGAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 479
Db 61 LeuLysSerMetThrAspAsnTyrLeuLeuAsnLeuAlaIleSerAspLeuLeuPheLeu 80
QY 480 CTCTACTGCTCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 539
Db 81 LeuThrValProPheThrPalaHisTyrAlaAlaGlnTyrPaspPheGlyAsnThrMet 100
QY 540 TGTCAACTCTTGACAGGCTCTATTTTATAGGCTTCTCTGGAATCTTCTTCATCATC 599
Db 101 CysGlnLeuLeuThrGlyLeuTyrPheIleGlyPheSerGlyIlePhePheIleIle 120
QY 600 CTCTGTCACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 659
Db 121 LeuLeuThrIleAspArgTyrLeuAlaIleValHisAlaValPheAlaLeuLysAlaArg 140
QY 660 ACGGTCACTTTGGGGTGGTGAAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 719
Db 141 ThrValThrPheGlyValValThrSerValIleThrTrpValValAlaValPheAlaSer 160
QY 720 CTCCAGGATCATCTTTACAGATGATCAAAAAGAGTCTTATACCTGCGAGCTCT 779
```


Db 161 LeuProGlyIleIlePheThrArgSerGlnArgGluGlyLeuHisTyrThrCysSerSer 180
QY 780 CATTTCATAC-----ATTAAAGATAGTCAT 806
|||||
Db 181 HisPheProTyrSerGlnTyrGlnPheTyrLysAsnPheGlnThrLeu-LysMetValI 200
|||
QY 807 CTTGGGCTGGTCTCGCTGCTGTCATGGTCATCTGCTACTCGGGAATCCTAAAC 866
|||||
Db 200 eLeuGlyLeuValProLeuLeuValMetValIleCysTyrSerGlyIleLeuLysTh 220
|||||
QY 867 TCTGCTTCGGTGTGGAATGAGAAAGAGGACAGGCTGTGAGGCTTATCTCACCAT 926
|||||
Db 220 rLeuLeuArgCysArgAsnGluLysArgHisArgAlaValArgLeuIlePheThrI 240
|||||
QY 927 CAGATGTTTATTTCTCTCTGGCTCCCTACACATGCTCTCTCTCTGACACCTT 986
|||||
Db 240 eMetIleValTyrPheLeuPheTrpAlaProTyrAsnIleValLeuLeuAsnThrPh 260
|||||
QY 987 CCAGGAATCTTTGGCTGAATAATTCAGTAGCTTAACAGGTGGACCAAGCTATGCA 1046
|||||
Db 260 eGlnGluPheGlyLeuAsnAsnCysSerCysAsnArgLeuAspIleAlaMetGI 280
|||||
QY 1047 GGTGACAGAGACTCTTGGGATGAGGCACTGCTGCATCAACCCCATCATCTATGCTTGT 1106
|||||
Db 280 nValThrGluThrLeuGlyMetThrHisCysCysIleAsnProIleIleTyrAlaPheVa 300
|||||
QY 1107 CCGGAGAGACTTCAGAACTACCTCTTAGTCTTCTTCCAAAGACATTCGCCAACGCTT 1166
|||||
Db 300 lGlyGluLysPheArgAsnTyrLeuLeuValPhePheGlnLysHisIleAlaLysArgPh 320
|||||
QY 1167 CTGCAATGCTGTCTATTTCACAGAGAGGCTCCGAGCGACGACAGCTCAGTTTACAC 1226
|||||
Db 320 eCysLysCysSerIlePheGlnGlnGluAlaProGluArgAlaSerSerValTyrTh 340
|||||
QY 1227 CCGATCCACTGGGAGACGAAATATCTGTGGGCTTG 1263
|||||
Db 340 rArgSerThrGlyGluGlnGluIleSerValGlyLeu 352
|||||

Search completed: June 3, 2003, 18:49:48
Job time : 119.268 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 3, 2003, 18:39:24 ; Search time 4.86523 Seconds
(without alignments)
5149.005 Million cell updates/sec

Title: US-09-938-703-3_COPY_792_885

Perfect score: 164

Sequences: 1 ATTAAGATGATCACTCTGG.....CTCTGCTTCGGTGTGGAAT 94

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 1816940

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame-n2p.model -DEV=xlp
-Q/cygn2_1/USPTO_spool/US09938703/runat_03062003_161406_22796/app_query.fasta_1.1870
-DB-A_Geneseq 101002 -QMT-fastan -SUFFIX=rag -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS-human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09938703.qcgn_1.162_@runat_03062003_161406_22796 -NCPU=6 -ICPU=3
-NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

A_Geneseq 101002.*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID | Description |
|------------|-------|---------------|--------|-------|--------------------|
| 1 | 164 | 100.0 | 215 | 18 | Inactive human CCR |
| 2 | 164 | 100.0 | 215 | 20 | HIV-1 co-receptor |
| 3 | 152 | 92.7 | 332 | 18 | Human chemokine re |
| 4 | 152 | 92.7 | 352 | 18 | Human CCR5, Homo |
| 5 | 152 | 92.7 | 352 | 18 | Human chemokine re |
| 6 | 152 | 92.7 | 352 | 18 | Human G-protein ch |
| 7 | 152 | 92.7 | 352 | 19 | Human CC chemokine |
| 8 | 152 | 92.7 | 352 | 20 | HIV-1 co-receptor |
| 9 | 152 | 92.7 | 352 | 21 | Human G-protein ch |
| 10 | 152 | 92.7 | 352 | 22 | Non-endogenous hum |
| 11 | 152 | 92.7 | 352 | 22 | Human CCR5 protein |
| 12 | 152 | 92.7 | 352 | 22 | Human HIV-1 co-rec |
| 13 | 152 | 92.7 | 352 | 22 | Amino acid sequenc |
| 14 | 152 | 92.7 | 352 | 22 | Human G-protein ch |
| 15 | 152 | 92.7 | 352 | 22 | Human G-protein ch |
| 16 | 152 | 92.7 | 352 | 22 | Human G-protein ch |
| 17 | 152 | 92.7 | 352 | 22 | Human CCR5 protein |
| 18 | 152 | 92.7 | 352 | 22 | Human chemokine re |
| 19 | 152 | 92.7 | 352 | 22 | Human HDGMR10 prot |
| 20 | 152 | 92.7 | 352 | 22 | Human G-protein ch |
| 21 | 152 | 92.7 | 352 | 23 | Human G-protein ch |
| 22 | 152 | 92.7 | 352 | 23 | Human G-protein ch |
| 23 | 152 | 92.7 | 352 | 23 | Human chemokine (C |
| 24 | 152 | 92.7 | 352 | 23 | Human CC chemokine |
| 25 | 152 | 92.7 | 352 | 23 | Human CCR5 Gln 55 |
| 26 | 152 | 92.7 | 371 | 19 | Human CC chemokine |
| 27 | 152 | 92.7 | 439 | 20 | Fusion protein con |
| 28 | 149 | 90.9 | 352 | 18 | Macaque chemokine |
| 29 | 134 | 81.7 | 329 | 22 | Human MCP-1 recept |
| 30 | 134 | 81.7 | 360 | 16 | Human monocyte che |
| 31 | 134 | 81.7 | 360 | 18 | Human monocyte che |
| 32 | 134 | 81.7 | 360 | 22 | Non-endogenous hum |
| 33 | 134 | 81.7 | 360 | 22 | Human CCR2b protei |
| 34 | 134 | 81.7 | 360 | 22 | Human CCR2-641 pol |
| 35 | 134 | 81.7 | 360 | 22 | Human wild-type CC |
| 36 | 134 | 81.7 | 374 | 16 | Human monocyte che |
| 37 | 134 | 81.7 | 374 | 22 | Human CCR2a protei |
| 38 | 131 | 79.9 | 354 | 19 | Mouse CC-CR5 prot |
| 39 | 127 | 77.4 | 331 | 19 | Peptide effecting |
| 40 | 109 | 66.5 | 332 | 23 | Human ovary specif |
| 41 | 109 | 66.5 | 355 | 17 | CC-chemokine recep |
| 42 | 109 | 66.5 | 355 | 17 | CC-chemokine recep |
| 43 | 109 | 66.5 | 355 | 17 | CC-chemokine recep |
| 44 | 109 | 66.5 | 355 | 18 | Human eosinophil e |
| 45 | 109 | 66.5 | 355 | 18 | Human chemokine re |

ALIGNMENTS

RESULT 1

AAW27408
ID AAW27408 standard; Protein; 215 AA.
XX
AC AAW27408;
XX
DT 14-APR-1998 (first entry)
XX
DE Inactive human CCR5.
XX

KW Inactive; human Cys-Cys chemokine receptor 5; CCR5;
KW human immunodeficiency virus; type 1; type 2; HIV-1; HIV-2;
KW predisposition; resistance; diagnosis; treatment; prevention;
KW inflammatory disease; rheumatoid arthritis; glomerulonephritis;
KW asthma; idiopathic pulmonary fibrosis; psoriasis; viral infection;
KW cancer; atherosclerosis; autoimmune disorder.
XX
OS Homo sapiens.
XX

Alignment Scores:

Pred. No.: 4.4e-13 Length: 352
 Score: 152.00 Matches: 31
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 92.68% Indels: 0
 DB: 18 Gaps: 0

US-09-938-703-3_COPY_792_885 (1-94) x AAW07602 (1-352)

QY 2 TTAAGATAGTCATCTGGGCTGCTCGCTGTCATGTCATGCTACTCG 61

Db 196 LeuYsIlleValIleuGlyLeuValLeuProLeuValMetValIleCysTyrSer 215

QY 62 GGAATCCATAAACTGCTGGGTGCGAAAT 94

Db 216 GlyIleLeuYsThrIleuLeuArgCysArgAsn 226

RESULT 7

AAW23835

ID AAW23835 standard; Protein: 352 AA.

XX AC AAW23835;

XX DT 08-JUN-1998 (first entry)

XX DE Human CC chemokine receptor 5 (CCR5).

XX CC chemokine receptor 5; CCR5; G-protein coupled receptor;
 KW human immunodeficiency virus; HIV; CD4; AIDS; therapy;
 XX transgenic animal.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Domain 29..55

FT /label= I

FT /note= "transmembrane domain"

FT Region 109..120

FT /note= "extracellular loop-1 (Claim 19)"

FT Domain 104..126

FT /label= III

FT /note= "transmembrane domain"

FT Domain 143..171

FT /label= IV

FT /note= "transmembrane domain"

FT Region 187..210

FT /note= "extracellular loop-2 (Claim 19)"

FT Domain 194..219

FT /label= V

FT /note= "transmembrane domain"

FT Domain 238..258

FT /label= VI

FT /note= "transmembrane domain"

FT Region 261..276

FT /note= "extracellular loop-3 (Claim 19)"

FT Domain 277..300

FT /label= VII

FT /note= "transmembrane domain"

XX WO9745543-A2.

XX 04-DEC-1997.

XX 28-MAY-1997; 97WO-US09586.

XX 28-MAY-1996; 96US-0018508.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Alkhatib G, Berger EA, Broder CC, Combadiere C;
 PI Feng Y, Kennedy PE, Murphy PM;

XX

DR WPI; 1998-032650/03.

XX N-PSDB; AAT76920.

PT CC chemokine receptor 5 polypeptide - used to inhibit membrane
 fusion between HIV and a target cell

XX Claim 68; Fig 1C; 70pp; English.

XX

This protein sequence comprises of a novel human macrophage-selective
 CC chemokine receptor that has been designated CCR5. The sequence
 was deduced from an isolated cDNA clone (see AAT76920). An Ala127Ileu
 variant (see W3840) of CCR5 was also identified. The susceptibility
 of human macrophages to HIV infection depends on cell surface
 expression of CD4 and CCR5. CCR5 is a member of the 7-transmembrane
 superfamily of G-protein coupled cell surface molecules. It plays
 an essential role in the membrane fusion step of infection by some
 HIV isolates. The establishment of stable, non-human cell lines
 and transgenic mammals having cells that coexpress human CD4 and
 CCR5 provides valuable tools for research of HIV infection.
 CC Antibodies that bind to CCR5, CCR5 variants, and CCR5-binding
 CC agents capable of blocking membrane fusion between HIV and target
 CC cells represent potential anti-HIV therapeutics for macrophage
 CC tropic strains of HIV.

XX SQ Sequence 352 AA;

Alignment Scores:

Pred. No.: 4.4e-13 Length: 352
 Score: 152.00 Matches: 31
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 92.68% Indels: 0
 DB: 19 Gaps: 0

US-09-938-703-3_COPY_792_885 (1-94) x AAW23835 (1-352)

QY 2 TTAAGATAGTCATCTGGGCTGCTCGCTGTCATGTCATGCTACTCG 61

Db 196 LeuYsIlleValIleuGlyLeuValLeuProLeuValMetValIleCysTyrSer 215

QY 62 GGAATCCATAAACTGCTGGGTGCGAAAT 94

Db 216 GlyIleLeuYsThrIleuLeuArgCysArgAsn 226

RESULT 8

AAW88232

ID AAW88232 standard; Protein: 352 AA.

XX AC AAW88232;

XX DT 15-MAR-1999 (first entry)

XX HIV-1 co-receptor CCR5.

HIV-1; CCR5; CCR5m303; co-receptor; infection; diagnosis; AIDS;
 gene therapy; human.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Domain 32..56

FT /note= "transmembrane domain 1"

FT Domain 67..87

FT /note= "transmembrane domain 2"

FT Misc-difference 101

FT /note= "corresponds to TGT (Cys) in wild-type CCR5,
 TGA (Stop) in CCR5m303"

FT Domain 103..124

FT /note= "transmembrane domain 3"

FT Domain 142..167

FT /note= "transmembrane domain 4"

FT Domain 200..223

XX
DE
XX
KW
KW
KW
KW
KW
XX
OS
XX
PN
XX
PD
XX
PF
PP
PR
PR
PA
PA
XX
PI
XX
DR
XX
PT
PT
PT
XX
PS
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC

Human CCR5 protein.

Chemokine; tumour diagnosis; colorectal; prostatic; organ rejection;
inflammation; autoimmune disease; metastasis; bronchial asthma; lupus;
chronic bowel inflammation; rheumatoid arthritis; cystostatic;
anti-inflammatory; antiasthmatic; immunosuppressive; dermatological;
antirheumatic; antiarthritic.

Homo sapiens.
WO200172830-A2.
04-OCT-2001.
02-APR-2001; 2001WO-EPO3708.
31-MAR-2000; 2000DE-1016013.
(IFEP-) IPF PHARM GMBH.
(FORS/) FORSSMANN U.
Forssmann W, Adermann K, Heitland A, Spodsborg N;
WPI; 2001-626256/72.
Diagnostic agent containing two or more receptor-specific ligands,
useful for detecting tumors, inflammation etc., also therapeutic use of
ligand inhibitors -
Disclosure; Page 10; 26pp; German.

This invention describes a novel diagnostic agent (A) comprising at least
two different ligands (I) for receptors (II) that are implicated in
disease. (A) are used for the diagnosis of tumors (especially colorectal
or prostatic), organ rejection, inflammation and autoimmune diseases.
Also inhibitors of (I) are used therapeutically against tumors (and their
metastases), inflammation (particularly bronchial asthma or chronic bowel
inflammation), or autoimmune diseases (rheumatoid arthritis or lupus),
where the (cardio)vascular, lymphatic, respiratory, nervous, digestive,
endocrine, motor or urogenital systems or skin are affected, and bone
marrow diseases. The products of the invention are chemokine derivatives
which have cytostatic, antiinflammatory, antirheumatic, antiasthmatic,
immunosuppressive, dermatological, antihematoic, antiarthritic.
Chemokines act on specific tumor and inflammatory cells through a
constellation of chemokine receptors (CR), which control migration and
proliferation of these cells. AAG80045-AAG80128 represent human chemokine
fragments used to illustrate the method of the invention.

Sequence 352 AA:

Alignment Scores:
Pred. No.: 4.4e-13 Length: 352
Score: 152.00 Matches: 31
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 92.68% Indels: 0
DB: 22 Gaps: 0

US-09-938-703-3_COPY_792_885 (1-94) x AAG80111 (1-352)

QY 2 TTAACATGACTCATTGGGTGTCCTGCCGCTGTTGCATCGTCACTGCTACTCG 61
|||||
Db 196 LeuLysilleValleLeuGlyLeuValleProLeuValMetValleCysTrSer 215
|||||

QY 62 GGAATCCTAAAACTCTGCTCGGNGTCGAAAT 94
|||||
Db 216 GlylleLeuLysThrLeuLeuargCysargAsn 226
|||||

RESULT 12
AAB82948
ID AAB82948 standard; Protein; 352 AA.
XX

XX
DE
XX
KW
KW
KW
KW
KW
XX
OS
XX
PN
XX
PD
XX
PF
PP
PR
PR
PA
PA
XX
PI
XX
DR
XX
PT
PT
PT
XX
PS
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC

Human CCR5 protein.

Chemokine; tumour diagnosis; colorectal; prostatic; organ rejection;
inflammation; autoimmune disease; metastasis; bronchial asthma; lupus;
chronic bowel inflammation; rheumatoid arthritis; cystostatic;
anti-inflammatory; antiasthmatic; immunosuppressive; dermatological;
antirheumatic; antiarthritic.

Homo sapiens.
WO200172830-A2.
04-OCT-2001.
02-APR-2001; 2001WO-EPO3708.
31-MAR-2000; 2000DE-1016013.
(IFEP-) IPF PHARM GMBH.
(FORS/) FORSSMANN U.
Forssmann W, Adermann K, Heitland A, Spodsborg N;
WPI; 2001-626256/72.
Diagnostic agent containing two or more receptor-specific ligands,
useful for detecting tumors, inflammation etc., also therapeutic use of
ligand inhibitors -
Disclosure; Page 10; 26pp; German.

This invention describes a novel diagnostic agent (A) comprising at least
two different ligands (I) for receptors (II) that are implicated in
disease. (A) are used for the diagnosis of tumors (especially colorectal
or prostatic), organ rejection, inflammation and autoimmune diseases.
Also inhibitors of (I) are used therapeutically against tumors (and their
metastases), inflammation (particularly bronchial asthma or chronic bowel
inflammation), or autoimmune diseases (rheumatoid arthritis or lupus),
where the (cardio)vascular, lymphatic, respiratory, nervous, digestive,
endocrine, motor or urogenital systems or skin are affected, and bone
marrow diseases. The products of the invention are chemokine derivatives
which have cytostatic, antiinflammatory, antirheumatic, antiasthmatic,
immunosuppressive, dermatological, antihematoic, antiarthritic.
Chemokines act on specific tumor and inflammatory cells through a
constellation of chemokine receptors (CR), which control migration and
proliferation of these cells. AAG80045-AAG80128 represent human chemokine
fragments used to illustrate the method of the invention.

Sequence 352 AA:

Alignment Scores:
Pred. No.: 4.4e-13 Length: 352
Score: 152.00 Matches: 31
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 92.68% Indels: 0
DB: 22 Gaps: 0

US-09-938-703-3_COPY_792_885 (1-94) x ABB56342 (1-352)

QY 2 TTAACATGACTCATTGGGTGTCCTGCCGCTGTTGCATCGTCACTGCTACTCG 61
|||||
Db 196 LeuLysilleValleLeuGlyLeuValleProLeuValMetValleCysTrSer 215
|||||

QY 62 GGAATCCTAAAACTCTGCTCGGNGTCGAAAT 94
|||||
Db 216 GlylleLeuLysThrLeuLeuargCysargAsn 226
|||||

RESULT 11
AAG80111
ID AAG80111 standard; Protein; 352 AA.
XX
AC
AAG80111;
XX
DT 17-JAN-2002 (first entry)
XX

XX
DE
XX
KW
KW
KW
KW
KW
XX
OS
XX
PN
XX
PD
XX
PF
PP
PR
PR
PA
PA
XX
PI
XX
DR
XX
PT
PT
PT
XX
PS
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC

Human CCR5 protein.

Chemokine; tumour diagnosis; colorectal; prostatic; organ rejection;
inflammation; autoimmune disease; metastasis; bronchial asthma; lupus;
chronic bowel inflammation; rheumatoid arthritis; cystostatic;
anti-inflammatory; antiasthmatic; immunosuppressive; dermatological;
antirheumatic; antiarthritic.

Homo sapiens.
WO200172830-A2.
04-OCT-2001.
02-APR-2001; 2001WO-EPO3708.
31-MAR-2000; 2000DE-1016013.
(IFEP-) IPF PHARM GMBH.
(FORS/) FORSSMANN U.
Forssmann W, Adermann K, Heitland A, Spodsborg N;
WPI; 2001-626256/72.
Diagnostic agent containing two or more receptor-specific ligands,
useful for detecting tumors, inflammation etc., also therapeutic use of
ligand inhibitors -
Disclosure; Page 10; 26pp; German.

This invention describes a novel diagnostic agent (A) comprising at least
two different ligands (I) for receptors (II) that are implicated in
disease. (A) are used for the diagnosis of tumors (especially colorectal
or prostatic), organ rejection, inflammation and autoimmune diseases.
Also inhibitors of (I) are used therapeutically against tumors (and their
metastases), inflammation (particularly bronchial asthma or chronic bowel
inflammation), or autoimmune diseases (rheumatoid arthritis or lupus),
where the (cardio)vascular, lymphatic, respiratory, nervous, digestive,
endocrine, motor or urogenital systems or skin are affected, and bone
marrow diseases. The products of the invention are chemokine derivatives
which have cytostatic, antiinflammatory, antirheumatic, antiasthmatic,
immunosuppressive, dermatological, antihematoic, antiarthritic.
Chemokines act on specific tumor and inflammatory cells through a
constellation of chemokine receptors (CR), which control migration and
proliferation of these cells. AAG80045-AAG80128 represent human chemokine
fragments used to illustrate the method of the invention.

Sequence 352 AA:

Alignment Scores:
Pred. No.: 4.4e-13 Length: 352
Score: 152.00 Matches: 31
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 92.68% Indels: 0
DB: 22 Gaps: 0

US-09-938-703-3_COPY_792_885 (1-94) x ABB56342 (1-352)

QY 2 TTAACATGACTCATTGGGTGTCCTGCCGCTGTTGCATCGTCACTGCTACTCG 61
|||||
Db 196 LeuLysilleValleLeuGlyLeuValleProLeuValMetValleCysTrSer 215
|||||

QY 62 GGAATCCTAAAACTCTGCTCGGNGTCGAAAT 94
|||||
Db 216 GlylleLeuLysThrLeuLeuargCysargAsn 226
|||||

RESULT 11
AAG80111
ID AAG80111 standard; Protein; 352 AA.
XX
AC
AAG80111;
XX
DT 17-JAN-2002 (first entry)
XX

XX
DE
XX
KW
KW
KW
KW
KW
XX
OS
XX
PN
XX
PD
XX
PF
PP
PR
PR
PA
PA
XX
PI
XX
DR
XX
PT
PT
PT
XX
PS
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC

Human CCR5 protein.

Chemokine; tumour diagnosis; colorectal; prostatic; organ rejection;
inflammation; autoimmune disease; metastasis; bronchial asthma; lupus;
chronic bowel inflammation; rheumatoid arthritis; cystostatic;
anti-inflammatory; antiasthmatic; immunosuppressive; dermatological;
antirheumatic; antiarthritic.

Homo sapiens.
WO200172830-A2.
04-OCT-2001.
02-APR-2001; 2001WO-EPO3708.
31-MAR-2000; 2000DE-1016013.
(IFEP-) IPF PHARM GMBH.
(FORS/) FORSSMANN U.
Forssmann W, Adermann K, Heitland A, Spodsborg N;
WPI; 2001-626256/72.
Diagnostic agent containing two or more receptor-specific ligands,
useful for detecting tumors, inflammation etc., also therapeutic use of
ligand inhibitors -
Disclosure; Page 10; 26pp; German.

This invention describes a novel diagnostic agent (A) comprising at least
two different ligands (I) for receptors (II) that are implicated in
disease. (A) are used for the diagnosis of tumors (especially colorectal
or prostatic), organ rejection, inflammation and autoimmune diseases.
Also inhibitors of (I) are used therapeutically against tumors (and their
metastases), inflammation (particularly bronchial asthma or chronic bowel
inflammation), or autoimmune diseases (rheumatoid arthritis or lupus),
where the (cardio)vascular, lymphatic, respiratory, nervous, digestive,
endocrine, motor or urogenital systems or skin are affected, and bone
marrow diseases. The products of the invention are chemokine derivatives
which have cytostatic, antiinflammatory, antirheumatic, antiasthmatic,
immunosuppressive, dermatological, antihematoic, antiarthritic.
Chemokines act on specific tumor and

QY 62 GGATCTCTAAACCTCTGCTCGTCTCGTCTCGAAT 94
Db 216 GlyIleLeuLysThrLeuLeuArgCysArgAsn 226
RESULT 14
AAE07037
ID AAE07037 standard; Protein; 352 AA.
AC AAE07037;
XX 16-OCT-2001 (first entry)
XX Human G-protein chemokine receptor (CCR5) HDGNR10 protein #1.
XX Human; G-protein chemokine receptor; CCR5; HDGNR10; inflammation; HIV;
KW human immunodeficiency virus; antimicrobial; vasodilator; vulnery;
KW cytostatic; immunosuppressive; nootropic; neuroprotective; gene therapy;
KW neurodegenerative disorder; Kaposi's sarcoma; autoimmune disease;
KW rheumatoid arthritis; cancer; breast; ovary; adrenal gland; bone marrow;
KW gastrointestinal tract; lung; liver; immune disorder; Addison's disease;
KW haemolytic anaemia; autoimmune thyroiditis; diabetes mellitus; allergy;
KW multiple sclerosis; ulcerative colitis; Crohn's disease; wound healing;
KW cardiovascular disorder; myocardial ischaemia.
XX Homo sapiens.
XX
XX
FH Key Location/Qualifiers
FT Domain 1..36
FT /label= Extracellular_domain
FT Domain 37..305
FT /label= Transmembrane_domain
FT Domain 37..58
FT /label= Transmembrane_domain
FT /note= "Segment 1"
FT Domain 59..67
FT /label= Intracellular_loop_1
FT Domain 68..88
FT /label= Transmembrane_domain
FT /note= "Segment 2"
FT Domain 89..102
FT /label= Extracellular_loop_1
FT Domain 103..124
FT /label= Transmembrane_domain
FT /note= "Segment 3"
FT Domain 125..141
FT /label= Intracellular_loop_2
FT Domain 142..166
FT /label= Transmembrane_domain
FT /note= "Segment 4"
FT Domain 167..195
FT /label= Extracellular_loop_2
FT Domain 196..223
FT /label= Transmembrane_domain
FT /note= "Segment 5"
FT Domain 224..235
FT /label= Intracellular_loop_3
FT Domain 236..260
FT /label= Transmembrane_domain
FT /note= "Segment 6"
FT Domain 261..274
FT /label= Extracellular_loop_3
FT Domain 287..305
FT /label= Transmembrane_domain
FT /note= "Segment 7"
FT Domain 306..352
FT /label= Intracellular_domain
XX
XX WO200158915-A2.
XX 16-AUG-2001.
XX
XX 09-FEB-2001; 2001WO-US04152.
XX

PR 09-FEB-2000; 2000US-0181258.
PR 09-MAR-2000; 2000US-0187999.
PR 22-SEP-2000; 2000US-0234336.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Roschke V, Li Y, Ruben SM;
XX WPI; 2001-488965/53.
XX N-PSDB; AAD13181.
XX Isolated nucleic acid encoding a human G-protein chemokine receptor
PT (CCR5) HDGNR10 polypeptide, useful for preventing or treating
PT autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative
PT disorders and neurodegenerative disorders -
XX Claim 102; Fig 1; 495pp; English.
XX The present sequence is human G-protein chemokine receptor (CCR5) HDGNR10
CC protein. CCR5 HDGNR10 cDNA is included in ATCC Deposit No: 97183. CCR5
CC HDGNR10 antibodies are useful for treating, preventing or ameliorating a
CC disease or disorder associated with inflammation, defective or aberrant
CC chemotaxis of immune cells, HIV infection (such as Pneumocystis carinii
CC pneumonia or Kaposi's sarcoma) or defective or aberrant T-cell antigen
CC presenting cell interaction. The disease or disorder may also be an
CC infectious disease (e.g. a viral infection such as an early stage HIV
CC infection, a cytomegalovirus infection, or a poxvirus infection), an
CC autoimmune disease (e.g. rheumatoid arthritis) or a neurodegenerative
CC disorder. The disease or disorder may be associated with aberrant CCR5
CC expression, lack of CCR5 function, aberrant CCR5 ligand expression, or
CC lack of CCR5 ligand function. CCR5 HDGNR10 protein is used as a food
CC additive or preservative to increase or decrease storage capabilities.
CC CCR5 HDGNR10 DNA are useful for chromosome identification and in gene
CC therapy. CCR5 HDGNR10 DNA, protein, antibodies, agonists and antagonists
CC are also useful in the diagnosis, treatment and prevention of cancer
CC (breast, ovary, adrenal gland, bone, bone marrow, gastrointestinal tract,
CC liver, lung, urogenital); immune disorders (Addison's disease, allergies,
CC autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus,
CC Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC colitis); cardiovascular disorders (myocardial ischaemia) and
XX wound healing.
XX SQ Sequence 352 AA;
Alignment Scores:
Pred. No.: 4.4e-13 Length: 352
Score: 152.00 Matches: 31
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 92.68% Indels: 0
DB: 22 Gaps: 0
US-09-938-703-3_COPY_792_885 (1-94) x AAE07037 (1-352)
QY 2 TTAAAGATAGTCATCTGGGCTGCTCGCTGCTGTCATGTCATCTGTCG 61
Db 196 LeuLysIleValIleLeuGlyLeuValLeuProLeuValMetValIleCysIyrSer 215
QY 62 GGATCTCTAAACCTCTGCTCGTCTCGTCTCGAAT 94
Db 216 GlyIleLeuLysThrLeuLeuArgCysArgAsn 226
RESULT 15
AAE07039
ID AAE07039 standard; Protein; 352 AA.
XX AC AAE07039;
XX 16-OCT-2001 (first entry)
XX Human G-protein chemokine receptor (CCR5) HDGNR10 protein #2.
XX Human; G-protein chemokine receptor; CCR5; HDGNR10; inflammation; HIV;
XX

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OK nucleic - protein search, using frame_plus_n2p model

Run on: June 3, 2003, 18:44:09 ; Search time 1.80534 Seconds
(without alignments)
3063.973 Million cell updates/sec

Title: US-09-938-703-3_COPY_792_885

Perfect score: 164
Sequence: 1 ATTAAGATGATGCTCTGG.....CTCGCTCGGTCGCAAT 94

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 525148

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_n2p_model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US0938703/runat_03062003_161408_22885/app_query.fasta_1.1870
-DB=Issued Patents AA -SUFFIX=rai -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human4.0.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFWT=prt -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US0938703_cgn_1_1_42 -runat_03062003_161408_22885 -NCP0=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSFLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS.COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Length | ID | Description |
|------------|-------|-------|--------|----|-------------------|
| 1 | 164 | 100.0 | 34 | 4 | US-08-833-752-13 |
| 2 | 164 | 100.0 | 215 | 4 | US-09-087-232A-17 |
| 3 | 164 | 100.0 | 215 | 4 | US-08-833-752-6 |
| 4 | 152 | 92.7 | 54 | 4 | US-08-833-752-11 |
| 5 | 152 | 92.7 | 352 | 3 | US-08-466-343D-2 |
| 6 | 152 | 92.7 | 352 | 4 | US-09-087-232A-13 |
| 7 | 152 | 92.7 | 352 | 4 | US-08-861-105-14 |
| 8 | 152 | 92.7 | 352 | 4 | US-08-575-967A-2 |
| 9 | 152 | 92.7 | 352 | 4 | US-09-045-583-52 |
| 10 | 152 | 92.7 | 352 | 4 | US-09-517-605-5 |
| 11 | 152 | 92.7 | 352 | 4 | US-09-534-185-52 |
| 12 | 152 | 92.7 | 352 | 4 | US-08-833-752-5 |

| | | | | | | |
|----|-----|------|-----|---|-------------------|-------------------|
| 13 | 134 | 81.7 | 344 | 3 | US-08-466-343D-9 | Sequence 9, Appl |
| 14 | 134 | 81.7 | 347 | 1 | US-08-461-244-3 | Sequence 3, Appl |
| 15 | 134 | 81.7 | 360 | 1 | US-08-450-393A-4 | Sequence 4, Appl |
| 16 | 134 | 81.7 | 360 | 4 | US-08-446-669-4 | Sequence 4, Appl |
| 17 | 134 | 81.7 | 360 | 4 | US-09-045-583-50 | Sequence 50, Appl |
| 18 | 134 | 81.7 | 360 | 4 | US-09-045-583-51 | Sequence 51, Appl |
| 19 | 134 | 81.7 | 360 | 4 | US-09-534-185-50 | Sequence 50, Appl |
| 20 | 134 | 81.7 | 360 | 4 | US-09-534-185-51 | Sequence 51, Appl |
| 21 | 134 | 81.7 | 360 | 4 | US-08-833-752-7 | Sequence 7, Appl |
| 22 | 134 | 81.7 | 360 | 5 | PCT-US95-00476-4 | Sequence 4, Appl |
| 23 | 134 | 81.7 | 374 | 1 | US-08-450-393A-2 | Sequence 2, Appl |
| 24 | 134 | 81.7 | 374 | 4 | US-08-446-669-2 | Sequence 2, Appl |
| 25 | 134 | 81.7 | 374 | 5 | PCT-US95-00476-2 | Sequence 2, Appl |
| 26 | 131 | 79.9 | 354 | 4 | US-08-724-984A-2 | Sequence 2, Appl |
| 27 | 115 | 70.1 | 24 | 4 | US-09-087-232A-21 | Sequence 21, Appl |
| 28 | 109 | 66.5 | 355 | 4 | US-08-575-967A-4 | Sequence 4, Appl |
| 29 | 109 | 66.5 | 355 | 4 | US-08-847-296B-1 | Sequence 1, Appl |
| 30 | 109 | 66.5 | 355 | 4 | US-09-045-583-54 | Sequence 54, Appl |
| 31 | 109 | 66.5 | 355 | 4 | US-09-534-185-54 | Sequence 54, Appl |
| 32 | 107 | 65.2 | 353 | 4 | US-08-833-752-8 | Sequence 8, Appl |
| 33 | 106 | 64.6 | 355 | 1 | US-08-012-988A-2 | Sequence 2, Appl |
| 34 | 106 | 64.6 | 355 | 1 | US-08-450-393A-5 | Sequence 5, Appl |
| 35 | 106 | 64.6 | 355 | 4 | US-08-446-669-5 | Sequence 5, Appl |
| 36 | 106 | 64.6 | 355 | 4 | US-09-239-938-1 | Sequence 1, Appl |
| 37 | 106 | 64.6 | 355 | 5 | PCT-US95-00476-5 | Sequence 5, Appl |
| 38 | 105 | 64.0 | 355 | 4 | US-09-045-583-53 | Sequence 53, Appl |
| 39 | 105 | 64.0 | 355 | 4 | US-09-534-185-53 | Sequence 53, Appl |
| 40 | 103 | 62.8 | 355 | 4 | US-08-833-752-9 | Sequence 9, Appl |
| 41 | 95 | 57.9 | 360 | 4 | US-08-875-573-20 | Sequence 20, Appl |
| 42 | 95 | 57.9 | 360 | 4 | US-09-232-878-2 | Sequence 2, Appl |
| 43 | 95 | 57.9 | 360 | 4 | US-09-045-583-55 | Sequence 55, Appl |
| 44 | 95 | 57.9 | 360 | 4 | US-09-534-185-55 | Sequence 55, Appl |
| 45 | 95 | 57.9 | 360 | 4 | US-08-833-752-10 | Sequence 10, Appl |

ALIGNMENTS

RESULT 1

US-08-833-752-13

; Sequence 13, Application US/08833752

; Patent No. 6448375

; GENERAL INFORMATION:

; APPLICANT: SAMSON, MICHEL

; APPLICANT: PARMENTIER, MARC

; APPLICANT: VASSART, GILBERT

; APPLICANT: LIBERT, FREDERICK

; TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR

; TITLE OF INVENTION: AND NUCLEIC ACID MOLECULES ENCODING SAID RECEPTOR

; NUMBER OF SEQUENCES: 17

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Knobbe, Martens, Olson & Bear

; STREET: 620 Newport Center Drive 16th Floor

; CITY: Newport Beach

; STATE: CA

; COUNTRY: U.S.A.

; ZIP: 92660

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/833,752

; FILING DATE: 9-APR-1997

; CLASSIFICATION: 536

; ATTORNEY/AGENT INFORMATION:

; NAME: Altman, Daniel E

; REGISTRATION NUMBER: 34,115

; REFERENCE/DOCKET NUMBER:

; INFORMATION FOR SEQ ID NO: 13:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 34 amino acids

; TYPE: amino acid


```
/ APPLICANT: SAMSON, MICHEL
/ APPLICANT: PARMENTIER, MARC
/ APPLICANT: VASSART, GILBERT
/ APPLICANT: LIBERT, FREDERICK
/ TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR
/ NUMBER OF SEQUENCES: 17
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Knobbe, Martens, Olson & Bear
/ STREET: 620 Newport Center Drive 16th Floor
/ CITY: Newport Beach
/ STATE: CA
/ COUNTRY: U.S.A.
/ ZIP: 92660
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ OPERATING SYSTEM: IBM PC compatible
/ SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/833,752
/ FILING DATE: 9-APR-1997
/ CLASSIFICATION: 536
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Altman, Daniel E.
/ REGISTRATION NUMBER: 34,115
/ REFERENCE/DOCKET NUMBER:
/ INFORMATION FOR SEQ ID NO: 11:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 54 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ US-08-833-752-11

Alignment Scores:
Pred. No.: 3,13e-14 Length: 54
Score: 152.00 Matches: 31
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 92.68% Indels: 0
DB: 4 Gaps: 0

US-09-938-703-3_COPY_792_885 (1-94) x US-08-833-752-11 (1-54)
QY 2 TTAAGATAGTCACTCTGGGCTGCTCGCGCTGCTGTCATGTCATGCTACTGCG 61
Db 20 LeuYslleValIleLeuGlyLeuValProLeuLeuValMetValIleCysTyrSer 39
QY 62 GGAATCCTAAAACTCTGCTGCTGCGTGCAGAAAT 94
Db 40 GlylleLeuYslThrLeuLeuArgCysArgAsn 50

RESULT 5
US-08-466-343D-2
; Sequence 2, Application US/08466343D
; Patent No. 6025154
; GENERAL INFORMATION:
; APPLICANT: LL, YI
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING HUMAN G-PROTEIN
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVE., NW, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/087,232A
; FILING DATE: 28 MAY 1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/048,057
; FILING DATE: 30 MAY 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: KOLE, LISA B.
; REGISTRATION NUMBER: 35,225
; REFERENCE/DOCKET NUMBER: AP 31115
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 408-2628
; TELEFAX: (212) 765-2519
```

```
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/466,343D
/ FILING DATE: 06-JUN-1995
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: STEFFE, ERIC K.
/ REGISTRATION NUMBER: 36,688
/ REFERENCE/DOCKET NUMBER: 1488.1150000/EKS/KLM
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (202) 371-2600
/ TELEFAX: (202) 371-2540
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 352 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-08-466-343D-2

Alignment Scores:
Pred. No.: 3,52e-14 Length: 352
Score: 152.00 Matches: 31
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 92.68% Indels: 0
DB: 3 Gaps: 0

US-09-938-703-3_COPY_792_885 (1-94) x US-08-466-343D-2 (1-352)
QY 2 TTAAGATAGTCACTCTGGGCTGCTCGCGCTGCTGTCATGTCATGCTACTGCG 61
Db 196 LeuYslleValIleLeuGlyLeuValProLeuLeuValMetValIleCysTyrSer 215
QY 62 GGAATCCTAAAACTCTGCTGCTGCGTGCAGAAAT 94
Db 216 GlylleLeuYslThrLeuLeuArgCysArgAsn 226

RESULT 6
US-09-087-232A-13
; Sequence 13, Application US/09087232A
; Patent No. 6153431
; GENERAL INFORMATION:
; APPLICANT: Quillent et al.
; TITLE OF INVENTION: HUMAN IMMUNODEFICIENCY VIRUS CO-RECEPTOR
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Baker & Botts, L.L.P. attn. Lisa Kole
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/087,232A
; FILING DATE: 28 MAY 1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/048,057
; FILING DATE: 30 MAY 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: KOLE, LISA B.
; REGISTRATION NUMBER: 35,225
; REFERENCE/DOCKET NUMBER: AP 31115
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 408-2628
; TELEFAX: (212) 765-2519
```


Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 92.68% Indels: 0
DB: 4 Gaps: 0

US-09-938-703-3_COPY_792_885 (1-94) x US-08-575-967A-2 (1-352)

QY 2 TTAAGATAGTCATCTGGGCGTGGTCTGCTCATGTCATGTCATGTCG 61
|||||
Db 196 LeuLysileValIleLeuGlyLeuValleuProLeuLeuValMetValIleCysTyrSer 215
|||||

QY 62 GGAATCCTAAAACTCTGCTCGGTGCGAAAT 94
|||||

Db 216 GlyIleLeuLysThrLeuLeuArgCysArgAsn 226
|||||

RESULT 9

US-09-045-583-52
; Sequence 52, Application US/09045583
; Patent No. 6287805
; GENERAL INFORMATION:
; APPLICANT: Graham, Gerard J. et al.
; TITLE OF INVENTION: No. 6287805el Molecules of the G Protein-Coupled
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/045,583
; FILING DATE: 20-MAR-98
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: MNI-044
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 52:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 352 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
US-09-045-583-52

Alignment Scores:
Pred. No.: 3.52e-14 Length: 352
Score: 152.00 Matches: 31
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 92.68% Indels: 0
DB: 4 Gaps: 0

US-09-938-703-3_COPY_792_885 (1-94) x US-09-045-583-52 (1-352)

QY 2 TTAAGATAGTCATCTGGGCGTGGTCTGCTCATGTCATGTCATGTCG 61
|||||
Db 196 LeuLysileValIleLeuGlyLeuValleuProLeuLeuValMetValIleCysTyrSer 215
|||||

QY 62 GGAATCCTAAAACTCTGCTCGGTGCGAAAT 94
|||||

Db 216 GlyIleLeuLysThrLeuLeuArgCysArgAsn 226
|||||

RESULT 10

US-09-517-605-5
; Sequence 5, Application US/09517605
; Patent No. 6391567
; GENERAL INFORMATION:
; APPLICANT: Littman, Dan R.
; APPLICANT: Kwon, Douglas S.
; APPLICANT: van Kooyk, Yvette
; APPLICANT: Geijtenbeck, Theo
; TITLE OF INVENTION: METHODS OF USING A FACILITATOR OF RETROVIRAL ENTRY INTO
; FILE REFERENCE: 1049-1-017
; CURRENT APPLICATION NUMBER: US/09/517,605
; CURRENT FILING DATE: 2000-03-02
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-517-605-5

Alignment Scores:
Pred. No.: 3.52e-14 Length: 352
Score: 152.00 Matches: 31
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 92.68% Indels: 0
DB: 4 Gaps: 0

US-09-938-703-3_COPY_792_885 (1-94) x US-09-517-605-5 (1-352)

QY 2 TTAAGATAGTCATCTGGGCGTGGTCTGCTCATGTCATGTCATGTCG 61
|||||
Db 196 LeuLysileValIleLeuGlyLeuValleuProLeuLeuValMetValIleCysTyrSer 215
|||||

QY 62 GGAATCCTAAAACTCTGCTCGGTGCGAAAT 94
|||||

Db 216 GlyIleLeuLysThrLeuLeuArgCysArgAsn 226
|||||

RESULT 11

US-09-534-185-52
; Sequence 52, Application US/09534185
; Patent No. 6403767
; GENERAL INFORMATION:
; APPLICANT: Graham, Gerard J. et al.
; TITLE OF INVENTION: No. 6403767el Molecules of the G Protein-Coupled
; Heptaheical Receptor Superfamily and Uses
; Therefor
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/534,185
; FILING DATE: 24-Mar-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/045,583
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:

```

Alignment Scores:
Pred. No.: 3.52e-14 Length: 352
Score: 152.00 Matches: 31
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 92.68% Indels: 0
DB: 4 Gaps: 0

US-09-938-703-3_COPY_792_885 (1-94) x US-08-833-752-5 (1-352)

QY 2 TTAACATAGTCACTTGGGCTGCTCGCTCGCTGCTGTCATGGPCATGCTGCTACTCG 161
Db 196 LeuLysIleValIleLeuGlyLeuValLeuProLeuMetValIleCysIyrSer 115

QY 62 GGAATCCTAAACACTGCTCGTGGTGCGAAAT 94
Db 216 GlyIleLeuLysThrLeuLeuArgCysArgAsn 226

RESULT 13
US-08-466-343D-9
; Sequence 9, Application US/08466343D
; Patent No. 6025154
; GENERAL INFORMATION:
; APPLICANT: LI, YI
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING HUMAN G-PROTEIN
; TITLE OF INVENTION: CHEMOKINE RECEPTOR HDGNR10 (AS AMENDED)
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVE., NW, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,343D
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: STEFFE, ERIC K.
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1488.1150000/EKS/KLM
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 344 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-466-343D-9

Alignment Scores:
Pred. No.: 1.41e-11 Length: 344
Score: 134.00 Matches: 26
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 96.30% Mismatches: 0
Query Match: 81.71% Indels: 0
DB: 3 Gaps: 0

US-09-938-703-3_COPY_792_885 (1-94) x US-08-466-343D-9 (1-344)

QY 14 ATCTGGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTG 73
Db 191 IleLeuGlyIleValIleLeuProLeuMetValIleCysIyrSerGlyIleLeu 215

```

QY 74 ACTCTGCTCGGTGTCGAAT 94
|||||
Db 211 ThrLeuLeuArgCysArgAsn 217

RESULT 14
US-08-461-244-3
; Sequence 3, Application US/08461244
; Patent No. 5776729
; GENERAL INFORMATION:
; APPLICANT: Soppet, Daniel R.
; APPLICANT: Yi, Li
; APPLICANT: Ruben, Steven M.
; APPLICANT: Rosen, Craig A.
; TITLE OF INVENTION: HUMAN G-PROTEIN RECEPTOR HGBR32
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN, CECCHI,
; ADDRESSEE: STUART & OLSTEIN
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,244
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferraro, Gregory D.
; REGISTRATION NUMBER: 36,134
; REFERENCE/DOCKET NUMBER: 325800-445
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 347 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-461-244-3

Alignment Scores:
Pred. No.: 1,41e-11 Length: 347
Score: 134.00 Matches: 26
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 96.30% Mismatches: 0
Query Match: 81.71% Indels: 0
DB: 1 Gaps: 0

US-09-938-703-3_COPY_792_885 (1-94) x US-08-461-244-3 (1-347)

QY 14 ACTCTGCTCGGTGTCGAAT 94
|||||
Db 195 ThrLeuLeuValLeuProLeuLeuMetValLeuCysTyrSerGlyLeuLeuLys 214

QY 74 ACTCTGCTCGGTGTCGAAT 94
|||||
Db 215 ThrLeuLeuArgCysArgAsn 221

RESULT 15
US-08-450-393A-4
; Sequence 4, Application US/08450393A
; Patent No. 5707815
; GENERAL INFORMATION:
; APPLICANT: Charo, Israel

; APPLICANT: Coughlin, Shaun
; TITLE OF INVENTION: MAMMALIAN MONOCYTE CHEMOATTRACTANT
; TITLE OF INVENTION: PROTEIN RECEPTORS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooley Godward Castro Huddleson & Tatum
; STREET: 5 Palo Alto Square
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94306-2155
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/450,393A
; FILING DATE: May 25, 1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Cserr, Luann
; REGISTRATION NUMBER: 31,822
; REFERENCE/DOCKET NUMBER: UCAL-237/02US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-843-5165
; TELEFAX: 415-8857-0663
; TELEX: 380816COOLEYPA
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 360 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-450-393A-4

Alignment Scores:
Pred. No.: 1,42e-11 Length: 360
Score: 134.00 Matches: 26
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 96.30% Mismatches: 0
Query Match: 81.71% Indels: 0
DB: 1 Gaps: 0

US-09-938-703-3_COPY_792_885 (1-94) x US-08-450-393A-4 (1-360)

QY 14 ACTCTGCTCGGTGTCGAAT 94
|||||
Db 208 ThrLeuLeuValLeuProLeuLeuMetValLeuCysTyrSerGlyLeuLeuLys 227

QY 74 ACTCTGCTCGGTGTCGAAT 94
|||||
Db 228 ThrLeuLeuArgCysArgAsn 234

Search completed: June 3, 2003, 18:52:48
Job time : 2.80534 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2003, 19:21:31 ; Search time 31.4634 Seconds
(without alignments)
201.057 Million cell updates/sec

Title: US-09-938-703-6

Perfect score: 1122

Sequence: 1 MDYQSSPIYDINITYTSEPC.....AACHGHLGLGNPKNSASVSK 215

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/prodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2_6/prodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2_6/prodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2_6/prodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2_6/prodata/1/iaa/PCUTUS_COMB.pep.*
- 6: /cgn2_6/prodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------------|
| 1 | 1122 | 100.0 | 215 | 4 | US-09-087-232A-17 |
| 2 | 1122 | 100.0 | 215 | 4 | US-08-833-752-6 |
| 3 | 958 | 85.4 | 184 | 4 | US-08-833-752-4 |
| 4 | 958 | 85.4 | 352 | 4 | US-09-087-232A-13 |
| 5 | 958 | 85.4 | 352 | 4 | US-08-851-105-14 |
| 6 | 958 | 85.4 | 352 | 4 | US-08-575-967A-2 |
| 7 | 958 | 85.4 | 352 | 4 | US-08-833-752-5 |
| 8 | 952 | 84.8 | 352 | 4 | US-09-045-583-52 |
| 9 | 952 | 84.8 | 352 | 4 | US-09-534-185-52 |
| 10 | 943 | 84.0 | 352 | 3 | US-08-466-343D-2 |
| 11 | 936 | 83.4 | 352 | 4 | US-09-517-605-5 |
| 12 | 776 | 69.2 | 354 | 4 | US-08-724-984A-2 |
| 13 | 694 | 61.9 | 344 | 3 | US-08-466-343D-9 |
| 14 | 694 | 61.9 | 347 | 1 | US-08-461-244-3 |
| 15 | 694 | 61.9 | 360 | 1 | US-08-450-393A-4 |
| 16 | 694 | 61.9 | 360 | 4 | US-08-446-669-4 |
| 17 | 694 | 61.9 | 360 | 4 | US-09-045-583-50 |
| 18 | 694 | 61.9 | 360 | 4 | US-09-534-185-50 |
| 19 | 694 | 61.9 | 360 | 5 | PCT-US95-00476-4 |
| 20 | 694 | 61.9 | 374 | 1 | US-08-450-393A-2 |
| 21 | 694 | 61.9 | 374 | 4 | US-08-446-669-2 |
| 22 | 694 | 61.9 | 374 | 5 | PCT-US95-00476-2 |
| 23 | 692 | 61.7 | 360 | 4 | US-08-833-752-7 |
| 24 | 688 | 61.3 | 360 | 4 | US-09-045-583-51 |
| 25 | 688 | 61.3 | 360 | 4 | US-09-534-185-51 |
| 26 | 598 | 53.3 | 355 | 1 | US-08-012-988A-2 |
| 27 | 598 | 53.3 | 355 | 1 | US-08-450-393A-5 |

28 598 53.3 355 4 US-08-446-669-5 Sequence 5, Appli
29 598 53.3 355 4 US-09-239-938-1 Sequence 1, Appli
30 598 53.3 355 5 PCT-US95-00476-5 Sequence 9, Appli
31 594 52.9 355 4 US-08-833-752-9 Sequence 5, Appli
32 562 50.1 355 4 US-09-045-583-53 Sequence 53, Appli
33 562 50.1 355 4 US-09-534-185-53 Sequence 20, Appli
34 547.5 48.8 360 4 US-08-875-573-20 Sequence 53, Appli
35 547.5 48.8 360 4 US-09-232-878-2 Sequence 2, Appli
36 547.5 48.8 360 4 US-09-045-583-55 Sequence 55, Appli
37 547.5 48.8 360 4 US-09-534-185-55 Sequence 15, Appli
38 525 46.8 100 4 US-09-087-232A-15 Sequence 10, Appli
39 515.5 45.9 360 4 US-08-833-752-10 Sequence 4, Appli
40 506.5 45.1 355 4 US-08-847-296B-1 Sequence 1, Appli
41 506.5 45.1 355 4 US-09-045-583-54 Sequence 54, Appli
42 506.5 45.1 355 4 US-09-534-185-54 Sequence 54, Appli
43 506.5 45.1 355 4 US-08-833-752-8 Sequence 8, Appli
44 502.5 44.8 355 4 US-09-087-232A-18 Sequence 18, Appli
45 448 39.9 87 4

ALIGNMENTS

RESULT 1

US-09-087-232A-17
; Sequence 17, Application US/09087232A
; Patent No. 6153431
; GENERAL INFORMATION:
; APPLICANT: Quillent et al.
; TITLE OF INVENTION: HUMAN IMMUNODEFICIENCY VIRUS CO-RECEPTOR
; TITLE OF INVENTION: VARIANTS ASSOCIATED WITH RESISTANCE TO VIRUS INFECTION.
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Baker & Botts, L.L.P. attn. Lisa Kole
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION NUMBER: US/09/087,232A
; FILING DATE: 28 MAY 1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/048,057
; FILING DATE: 30 MAY 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: KOLE, LISA B.
; REFERENCE/DOCKET NUMBER: AP 31115
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 408-2628
; TELEFAX: (212) 765-2519
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 215 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-087-232A-17

Query Match 100.0%; Score 1122; DB 4; Length 215;
Best Local Similarity 100.0%; Pred. No. 6.4e-102;
Matches 215; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDYQSSPIYDINITYTSEPCQKINVKQIAARLLPPLYSLVFFGFGVGNMVLILLNCKR 60
|||||
Db 1 MDYQSSPIYDINITYTSEPCQKINVKQIAARLLPPLYSLVFFGFGVGNMVLILLNCKR 60
|||||

QY 61 LKSMTDIYLLNLAISDLFLFTVPFWAHYAAQWDFGNTMCQLLTGLYFGFFSGIFII 120
DB 61 LKSMTDIYLLNLAISDLFLFTVPFWAHYAAQWDFGNTMCQLLTGLYFGFFSGIFII 120
QY 121 LITDRIYLVAVHAFKARTVTFGVTSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
DB 121 LITDRIYLVAVHAFKARTVTFGVTSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
QY 181 HPPYIKOSHLAGPAAACHGHLILGNPKNSASYSK 215
DB 181 HPPYIKOSHLAGPAAACHGHLILGNPKNSASYSK 215

RESULT 2

US-08-833-752-6
; Sequence 6, Application US/08833752
; Patent No. 6448375
; GENERAL INFORMATION:
; APPLICANT: SAMSON, MICHEL
; APPLICANT: PARMENTIER, MARC
; APPLICANT: VASSART, GILBERT
; APPLICANT: LIBERT, FREDERICK
; TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR
; TITLE OF INVENTION: AND NUCLEIC ACID MOLECULES ENCODING SAID RECEPTOR
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/833,752
; FILING DATE: 9-APR-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Altman, Daniel E
; REGISTRATION NUMBER: 34,115
; REFERENCE/DOCKET NUMBER:
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 215 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-833-752-6

Query Match 100.0%; Score 1122; DB 4; Length 215;
Best Local Similarity 100.0%; Pred. No. 6.4e-102;
Matches 215; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDQVSSPIYDINYYTSEPCKINVKQIAARLLPPLYSLVFIFGVGNMLVILINCKR 60
DB 1 MDQVSSPIYDINYYTSEPCKINVKQIAARLLPPLYSLVFIFGVGNMLVILINCKR 60
QY 61 LKSMTDIYLLNLAISDLFLFTVPFWAHYAAQWDFGNTMCQLLTGLYFGFFSGIFII 120
DB 61 LKSMTDIYLLNLAISDLFLFTVPFWAHYAAQWDFGNTMCQLLTGLYFGFFSGIFII 120
QY 121 LITDRIYLVAVHAFKARTVTFGVTSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
DB 121 LITDRIYLVAVHAFKARTVTFGVTSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
QY 181 HPPYIKOSHLAGPAAACHGHLILGNPKNSASYSK 215
DB 181 HPPYIKOSHLAGPAAACHGHLILGNPKNSASYSK 215

RESULT 3

US-08-833-752-4
; Sequence 4, Application US/08833752
; Patent No. 6448375
; GENERAL INFORMATION:
; APPLICANT: SAMSON, MICHEL
; APPLICANT: PARMENTIER, MARC
; APPLICANT: VASSART, GILBERT
; APPLICANT: LIBERT, FREDERICK
; TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR
; TITLE OF INVENTION: AND NUCLEIC ACID MOLECULES ENCODING SAID RECEPTOR
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/833,752
; FILING DATE: 9-APR-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Altman, Daniel E
; REGISTRATION NUMBER: 34,115
; REFERENCE/DOCKET NUMBER:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 184 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-833-752-4

Query Match 85.4%; Score 958; DB 4; Length 184;
Best Local Similarity 100.0%; Pred. No. 4.9e-86;
Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDQVSSPIYDINYYTSEPCKINVKQIAARLLPPLYSLVFIFGVGNMLVILINCKR 60
DB 1 MDQVSSPIYDINYYTSEPCKINVKQIAARLLPPLYSLVFIFGVGNMLVILINCKR 60
QY 61 LKSMTDIYLLNLAISDLFLFTVPFWAHYAAQWDFGNTMCQLLTGLYFGFFSGIFII 120
DB 61 LKSMTDIYLLNLAISDLFLFTVPFWAHYAAQWDFGNTMCQLLTGLYFGFFSGIFII 120
QY 121 LITDRIYLVAVHAFKARTVTFGVTSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
DB 121 LITDRIYLVAVHAFKARTVTFGVTSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
QY 181 HPPY 184
DB 181 HPPY 184

RESULT 4

US-09-087-232A-13
; Sequence 13, Application US/09087232A
; Patent No. 6153431
; GENERAL INFORMATION:
; APPLICANT: Quillent et al.
; TITLE OF INVENTION: HUMAN IMMUNODEFICIENCY VIRUS CO-RECEPTOR
; TITLE OF INVENTION: VARIANTS ASSOCIATED WITH RESISTANCE TO VIRUS INFECTION.
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:

RESULT 5
US-08-861-105-14
; Sequence 14, Application US/08861105
; Patent No. 6258527
; GENERAL INFORMATION:
; APPLICANT: LITWMAN, DAN R.
; APPLICANT: DENG, HONGKUI
; APPLICANT: ELLMEIER, WILFRIED
; APPLICANT: LANDAU, NATHANIEL R.
; APPLICANT: LIU, RONG
; TITLE OF INVENTION: G-COUPLED RECEPTORS ASSOCIATED WITH
; TITLE OF INVENTION: MACROPHAGE-TROPIC HIV, AND DIAGNOSTIC AND THERAPEUTIC
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David A. Jackson, Esq.
; STREET: 411 Hackensack Ave, Continental Plaza, 4th
; FLOOR
; STREET: Floor
; CITY: Hackensack

RESULT 6
US-08-575-967A-2
; Sequence 2, Application US/08575967A
; Patent No. 6285184
; GENERAL INFORMATION:
; APPLICANT: Gray et al.
; TITLE OF INVENTION: Chemokine Receptor Materials and Methods
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 S. Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA

ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/575,967A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: No. 6265184and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 32918
TELEPHONE: 206-485-1900
TELEFAX: 206-485-1662
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 352 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: /- "88C amino acid sequence"

Query Match 85.4%; Score 958; DB 4; Length 352;
Best Local Similarity 100.0%; Pred. No. 1e-85;
Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDYQSSPIYDINTYTSEPCQKINVKQIAARLLPPLYSILVTFEGVGNMLVILLINCKR 60
|||||
Db 1 MDYQSSPIYDINTYTSEPCQKINVKQIAARLLPPLYSILVTFEGVGNMLVILLINCKR 60
QY 61 LKSMTDIYLLNLAIISDLFFLLTPFPWAHYAAQWDFGNTMCQLTGLYFGFFSGIFFII 120
|||||
Db 61 LKSMTDIYLLNLAIISDLFFLLTPFPWAHYAAQWDFGNTMCQLTGLYFGFFSGIFFII 120
QY 121 LITDRIYLVAVHAFKARTVFGVTVSVITWVAVFASLPGLIIFTRSKQGLHYTCSS 180
|||||
Db 121 LITDRIYLVAVHAFKARTVFGVTVSVITWVAVFASLPGLIIFTRSKQGLHYTCSS 180
QY 181 HPFY 184
|||||
Db 181 HPFY 184

RESULT 7
US-08-833-752-5
Sequence 5, Application US/08833752
Patent No. 6448375
GENERAL INFORMATION:
APPLICANT: SAMSON, MICHEL
APPLICANT: PARMENTIER, MARC
APPLICANT: VASSART, GILBERT
APPLICANT: LIBERT, FREDERICK
TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR
TITLE OF INVENTION: AND NUCLEIC ACID MOLECULES ENCODING SAID RECEPTOR
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe, Martens, Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: CA
COUNTRY: U.S.A.
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25 (EPO)

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/833,752
FILING DATE: 9-APR-1997
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Altman, Daniel E.
REGISTRATION NUMBER: 34,115
REFERENCE/DOCKET NUMBER:
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 352 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-833-752-5
Query Match 85.4%; Score 958; DB 4; Length 352;
Best Local Similarity 100.0%; Pred. No. 1e-85;
Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDYQSSPIYDINTYTSEPCQKINVKQIAARLLPPLYSILVTFEGVGNMLVILLINCKR 60
|||||
Db 1 MDYQSSPIYDINTYTSEPCQKINVKQIAARLLPPLYSILVTFEGVGNMLVILLINCKR 60
QY 61 LKSMTDIYLLNLAIISDLFFLLTPFPWAHYAAQWDFGNTMCQLTGLYFGFFSGIFFII 120
|||||
Db 61 LKSMTDIYLLNLAIISDLFFLLTPFPWAHYAAQWDFGNTMCQLTGLYFGFFSGIFFII 120
QY 121 LITDRIYLVAVHAFKARTVFGVTVSVITWVAVFASLPGLIIFTRSKQGLHYTCSS 180
|||||
Db 121 LITDRIYLVAVHAFKARTVFGVTVSVITWVAVFASLPGLIIFTRSKQGLHYTCSS 180
QY 181 HPFY 184
|||||
Db 181 HPFY 184

RESULT 8
US-09-045-583-52
Sequence 52, Application US/09045583
Patent No. 6287805
GENERAL INFORMATION:
APPLICANT: Graham, Gerard J. et al.
TITLE OF INVENTION: No. 6287805el Molecules of the G Protein-Coupled
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD, LLP
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/045,583
FILING DATE: 20-MAR-98
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Mandragouras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: ANI-044
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 52:
SEQUENCE CHARACTERISTICS:

;
; LENGTH: 352 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
US-09-045-583-52

Query Match 84.8%; Score 952; DB 4; Length 352;
Best Local Similarity 98.9%; Pred. No. 3.9e-85;
Matches 182; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDYQSSPIYDINITYTSEPCQKINVKQIAARLLPPLYSLVIFGVGNMVLILINCKR 60
DB 1 MDYQSSPIYDIDYITSEPCQKINVKQIAARLLPPLYSLVIFGVGNMVLILINCKR 60
QY 61 LKSMTDIYLLNLAIISDLFFLLTPFWAHYAAAQWDFGNTMCOLLGLYFIFGFFSGIFFII 120
DB 61 LKSMTDIYLLNLAIISDLFFLLTPFWAHYAAAQWDFGNTMCOLLGLYFIFGFFSGIFFII 120
QY 121 LTIIDRYLAVVHAFKARTVTGVTSTVITWVAVFASLPGIIFTRSQKGLHYTCSS 180
DB 121 LTIIDRYLAVVHAFKARTVTGVTSTVITWVAVFASLPGIIFTRSQKGLHYTCSS 180
QY 181 HFPY 184
DB 181 HFPY 184

RESULT 9
US-09-534-185-52
; Sequence 52, Application US/09534185
; Patent No. 6403767
; GENERAL INFORMATION:
; APPLICANT: Graham, Gerard J. et al.
; TITLE OF INVENTION: No. 6403767el Molecules of the G Protein-Coupled
; Heptahelical Receptor Superfamily and Uses
; Therefor

NUMBER OF SEQUENCES: 56
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD, LLP
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/534,185
FILING DATE: 24-Mar-2000
CLASSIFICATION: <Unknown>
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 09/045,583
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Mandragouras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: MNI-044
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 52:
SEQUENCE CHARACTERISTICS:
LENGTH: 352 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
SEQUENCE DESCRIPTION: SEQ ID NO: 52:
US-09-534-185-52

Query Match 84.8%; Score 952; DB 4; Length 352;
Best Local Similarity 98.9%; Pred. No. 3.9e-85;
Matches 182; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDYQSSPIYDINITYTSEPCQKINVKQIAARLLPPLYSLVIFGVGNMVLILINCKR 60
DB 1 MDYQSSPIYDIDYITSEPCQKINVKQIAARLLPPLYSLVIFGVGNMVLILINCKR 60
QY 61 LKSMTDIYLLNLAIISDLFFLLTPFWAHYAAAQWDFGNTMCOLLGLYFIFGFFSGIFFII 120
DB 61 LKSMTDIYLLNLAIISDLFFLLTPFWAHYAAAQWDFGNTMCOLLGLYFIFGFFSGIFFII 120
QY 121 LTIIDRYLAVVHAFKARTVTGVTSTVITWVAVFASLPGIIFTRSQKGLHYTCSS 180
DB 121 LTIIDRYLAVVHAFKARTVTGVTSTVITWVAVFASLPGIIFTRSQKGLHYTCSS 180
QY 181 HFPY 184
DB 181 HFPY 184

RESULT 10
US-08-466-343D-2
; Sequence 2, Application US/08466343D
; Patent No. 6025154
; GENERAL INFORMATION:
; APPLICANT: LI, YI
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING HUMAN G-PROTEIN
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVE., NW, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/466,343D
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: STEFFE, ERIC K.
REGISTRATION NUMBER: 36,688
REFERENCE/DOCKET NUMBER: 1488.1150000/EKS/KLM
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 352 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-466-343D-2

Query Match 84.0%; Score 943; DB 3; Length 352;
Best Local Similarity 97.8%; Pred. No. 2.9e-84;
Matches 180; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 1 MDYQSSPIYDINITYTSEPCQKINVKQIAARLLPPLYSLVIFGVGNMVLILINCKR 60
DB 1 MDYQSSPIYDINITYTSEPCQKINVKQIAARLLPPLYSLVIFGVGNMVLILINCKR 60
QY 61 LKSMTDIYLLNLAIISDLFFLLTPFWAHYAAAQWDFGNTMCOLLGLYFIFGFFSGIFFII 120
DB 61 LKSMTDIYLLNLAIISDLFFLLTPFWAHYAAAQWDFGNTMCOLLGLYFIFGFFSGIFFII 120


```

RESULT 12
US-08-724-984A-2
; Sequence 2, Application US/08724984A
; Patent No. 6388055
; GENERAL INFORMATION:
; APPLICANT: Derk Bergsma, Mary Brawner, and Usman Shabon
; TITLE OF INVENTION: No. 6388055el Mouse Genomic Clone of the CC-
; TITLE OF INVENTION: CKRS Receptor
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road, P.O. Box 1539
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
; COMPUTER: IBM 486
; OPERATING SYSTEM: WINDOWS FOR WORKGROUPS

```

RESULT 13
US-08-466-343D-9
Sequence 9, Application US/08466343D
Patent No. 6025154
GENERAL INFORMATION:
APPLICANT: LI, Yi
TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING HUMAN G-PROTEIN
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSES: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
STREET: 1100 NEW YORK AVE., NW, SUITE 600
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/466,343D
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: STEFFE, ERIC K.
REGISTRATION NUMBER: 36,688
REFERENCE/DOCKET NUMBER: 1488.1150000/EKS/KLM
TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 371-2600
 TELEFAX: (202) 371-2540
 INFORMATION FOR SEQ ID NO: 9:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 344 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-466-343D-9

Query Match 51.9%; Score 694; DB 3; Length 344;
 Best Local Similarity 75.9%; Pred. No. 4.9e-60;
 Matches 132; Conservative 17; Mismatches 23; Indels 2; Gaps 1;

QY 10 YDINYTTSEPCOKINVKQIAARLLPPLYSLVFIFGVGNMLVILLINCKRLKSMTDIYL 69
 Db 7 FDYDY--GAPCHKFDVKQIGAQQLPPLYSLVFIFGVGNMLVILLINCKRLKCLTDIYL 64
 QY 70 LNLASDLFFLLTPFWAHYAAQWDFGNTMCCOLLTGLYFGFSGIFFIILLTIDRYLA 129
 Db 65 LNLASDLFFLLTPFWAHYAAQWDFGNTMCCOLLTGLYFGFSGIFFIILLTIDRYLA 124
 QY 130 VHAVFALKARTVTFGVVTSVITWVAVFASLPGLIFTRSQEGLHYTCSSHP 183
 Db 125 IVHAVFALKARTVTFGVVTSVITWVAVFASLPGLIFTRSQEGLHYTCSSHP 178

RESULT 14

US-08-461-244-3
 Sequence 3, Application US/08461244
 Patent No. 5776729

GENERAL INFORMATION:
 APPLICANT: Soppet, Daniel R.
 APPLICANT: Yi, Li
 APPLICANT: Ruben, Steven M.
 APPLICANT: Rosen, Craig A.
 TITLE OF INVENTION: HUMAN G-PROTEIN RECEPTOR HGBR32
 NUMBER OF SEQUENCES: 7
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN, CECCHI,
 ADDRESSEE: STUART & OLSTEIN
 STREET: 6 Becker Farm Road
 CITY: Roseland
 STATE: New Jersey
 COUNTRY: USA
 ZIP: 07068
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/461,244
 FILING DATE: 05-JUN-1995
 CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:
 NAME: Ferraro, Gregory D.
 REGISTRATION NUMBER: 36,134
 REFERENCE/DOCKET NUMBER: 325800-445
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 201-994-1700
 TELEFAX: 201-994-1744
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 347 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-461-244-3

Query Match 61.9%; Score 694; DB 1; Length 347;

Best Local Similarity 75.9%; Pred. No. 4.9e-60;
 Matches 132; Conservative 17; Mismatches 23; Indels 2; Gaps 1;

QY 10 YDINYTTSEPCOKINVKQIAARLLPPLYSLVFIFGVGNMLVILLINCKRLKSMTDIYL 69
 Db 11 FDYDY--GAPCHKFDVKQIGAQQLPPLYSLVFIFGVGNMLVILLINCKRLKCLTDIYL 68
 QY 70 LNLASDLFFLLTPFWAHYAAQWDFGNTMCCOLLTGLYFGFSGIFFIILLTIDRYLA 129
 Db 69 LNLASDLFFLLTPFWAHYAAQWDFGNTMCCOLLTGLYFGFSGIFFIILLTIDRYLA 128
 QY 130 VHAVFALKARTVTFGVVTSVITWVAVFASLPGLIFTRSQEGLHYTCSSHP 183
 Db 129 IVHAVFALKARTVTFGVVTSVITWVAVFASLPGLIFTRSQEGLHYTCSSHP 182

RESULT 15

US-08-450-393A-4
 Sequence 4, Application US/08450393A
 Patent No. 5707815

GENERAL INFORMATION:
 APPLICANT: Charo, Israel
 APPLICANT: Coughlin, Shaun
 TITLE OF INVENTION: MAMMALIAN MONOCYTE CHERMOATTRACTANT
 NUMBER OF SEQUENCES: 14
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Cooley Godward Castro Huddleson & Tatum
 STREET: 5 Palo Alto Square
 CITY: Palo Alto
 STATE: California
 COUNTRY: USA
 ZIP: 94306-2155
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/450,393A
 FILING DATE: May 25, 1995
 CLASSIFICATION: 424
 ATTORNEY/AGENT INFORMATION:
 NAME: Cseri, Luann
 REGISTRATION NUMBER: 31,822
 REFERENCE/DOCKET NUMBER: UCAL-237/0205
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-843-5165
 TELEFAX: 415-8857-0663
 TELEX: 380816CooleyPA
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 360 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-450-393A-4

Query Match 61.9%; Score 694; DB 1; Length 360;

Best Local Similarity 75.9%; Pred. No. 5.1e-60;
 Matches 132; Conservative 17; Mismatches 23; Indels 2; Gaps 1;

QY 10 YDINYTTSEPCOKINVKQIAARLLPPLYSLVFIFGVGNMLVILLINCKRLKSMTDIYL 69
 Db 24 FDYDY--GAPCHKFDVKQIGAQQLPPLYSLVFIFGVGNMLVILLINCKRLKCLTDIYL 81
 QY 70 LNLASDLFFLLTPFWAHYAAQWDFGNTMCCOLLTGLYFGFSGIFFIILLTIDRYLA 129
 Db 82 LNLASDLFFLLTPFWAHYAAQWDFGNTMCCOLLTGLYFGFSGIFFIILLTIDRYLA 141
 QY 130 VHAVFALKARTVTFGVVTSVITWVAVFASLPGLIFTRSQEGLHYTCSSHP 183
 Db 142 IVHAVFALKARTVTFGVVTSVITWVAVFASLPGLIFTRSQEGLHYTCSSHP 195

Search completed: June 3, 2003, 19:25:10
Job time : 32.4634 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2003, 19:17:26 ; Search time 55.061 Seconds
(without alignments)
520.312 Million cell updates/sec

Title: US-09-938-703-6

Perfect score: 1122

Sequence: 1 MDQVSSPIVDNYTSEPC.....AACHGHLILGNPKNSASVSK 215

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

A_Geneseq_101002.*

1: /SID2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|-------------|--------------------|
| 1 | 1122 | 100.0 | 215 | 18 AAW27408 | Inactive human CCR |
| 2 | 1122 | 100.0 | 215 | 20 AAW8238 | HIV-1 co-receptor |
| 3 | 958 | 85.4 | 184 | 18 AAW27406 | Inactive human CCR |
| 4 | 958 | 85.4 | 352 | 18 AAW27407 | Human CCR5, Homo |
| 5 | 958 | 85.4 | 352 | 18 AAW27123 | Human chemokine re |
| 6 | 958 | 85.4 | 352 | 19 AAW23835 | Human CC chemokine |
| 7 | 958 | 85.4 | 352 | 20 AAW8232 | HIV-1 co-receptor |
| 8 | 958 | 85.4 | 352 | 22 AAW56342 | Non-endogenous hum |
| 9 | 958 | 85.4 | 352 | 22 AAW80111 | Human CCR5 protein |
| 10 | 958 | 85.4 | 352 | 22 AAW82948 | Human HIV-1 co-rec |

| | | | | | |
|----|-------|------|-----|-------------|--------------------|
| 11 | 958 | 85.4 | 352 | 22 AAE07039 | Human G-protein ch |
| 12 | 958 | 85.4 | 352 | 22 AAE07048 | Human G-protein ch |
| 13 | 958 | 85.4 | 352 | 22 AAB83354 | Human CCR5 protein |
| 14 | 958 | 85.4 | 352 | 22 AAE04321 | Human chemokine re |
| 15 | 958 | 85.4 | 352 | 22 AAB46858 | Human HDGMR10 prot |
| 16 | 958 | 85.4 | 352 | 23 AA097152 | Human G-protein ch |
| 17 | 958 | 85.4 | 352 | 23 ABB08343 | Human chemokine (C |
| 18 | 958 | 85.4 | 352 | 23 AAM52828 | Human CC chemokine |
| 19 | 958 | 85.4 | 439 | 20 AAY41280 | Fusion protein con |
| 20 | 953 | 84.9 | 371 | 19 AAW23834 | Human CC chemokine |
| 21 | 952 | 84.8 | 352 | 23 AAW52829 | Human CCR5 Gln 55 |
| 22 | 943 | 84.0 | 352 | 18 AAW07602 | Human G-protein ch |
| 23 | 943 | 84.0 | 352 | 21 AAY80128 | Human G-protein ch |
| 24 | 943 | 84.0 | 352 | 22 AAE07037 | Human G-protein ch |
| 25 | 943 | 84.0 | 352 | 22 AAE07046 | Human G-protein ch |
| 26 | 943 | 84.0 | 352 | 23 AAU97150 | Human G-protein ch |
| 27 | 936 | 83.4 | 352 | 22 AAG79089 | Amino acid sequenc |
| 28 | 933 | 83.2 | 352 | 18 AAW27125 | Macaque chemokine |
| 29 | 839 | 74.8 | 332 | 18 AAW26766 | Human chemokine re |
| 30 | 776 | 69.2 | 354 | 19 AAW54037 | Mouse CC-CKR5 prot |
| 31 | 695 | 61.9 | 360 | 22 AAU07613 | Human CCR2-641 pol |
| 32 | 694 | 61.9 | 360 | 16 AAR79166 | Human monocyte che |
| 33 | 694 | 61.9 | 360 | 18 AAW35833 | Human monocyte che |
| 34 | 694 | 61.9 | 360 | 22 ABB56340 | Non-endogenous hum |
| 35 | 694 | 61.9 | 360 | 22 AAG80108 | Human CCR2b protei |
| 36 | 694 | 61.9 | 360 | 22 AAU07614 | Human wild-type CC |
| 37 | 694 | 61.9 | 374 | 16 AAR79165 | Human monocyte che |
| 38 | 694 | 61.9 | 374 | 22 AAG80107 | Human CCR2a protei |
| 39 | 608.5 | 54.2 | 329 | 22 AAB46859 | Human MCP-1 recept |
| 40 | 598 | 53.3 | 355 | 15 AAR52749 | C-C chemokine rece |
| 41 | 598 | 53.3 | 355 | 18 AAW26588 | Human MIP-1 alpha/ |
| 42 | 598 | 53.3 | 355 | 18 AAW25751 | Human MIP-1alpha/R |
| 43 | 598 | 53.3 | 355 | 21 AAB20571 | Human CC-chemokine |
| 44 | 565.5 | 50.4 | 355 | 18 AAW29179 | Rat CC chemokine r |
| 45 | 547.5 | 48.8 | 360 | 17 AAR99274 | Chemokine receptor |

ALIGNMENTS

RESULT 1

AAW27408
ID AAW27408 standard; Protein; 215 AA.

XX AC AAW27408;

DT 14-APR-1998 (first entry)

DE Inactive human CCR5.

XX KW Inactive; human Cys-Cys chemokine receptor 5; CCR5;
KW human immunodeficiency virus; type 1; type 2; HIV-1; HIV-2;
KW predisposition; resistance; diagnosis; treatment; prevention;
KW inflammatory disease; rheumatoid arthritis; glomerulonephritis;
KW asthma; idiopathic pulmonary fibrosis; psoriasis; viral infection;
KW cancer; atherosclerosis; autoimmune disorder.

XX OS Homo sapiens.

XX PN WO9732019-A2.

XX PD 04-SEP-1997.

XX PF 28-FEB-1997; 97WO-BE00023.

XX PR 06-AUG-1996; 96EP-0870102.

XX PR 01-MAR-1996; 96EP-0870021.

XX PA (EURO-) EUROSREEN SA.

XX PI Libert F, Parmentier M, Samson M, Vassart G;

XX DR WPI; 1997-479829/44.

```

DR N-PSDB; AAT90118.
XX
PT Active and inactive forms of human CC chemokine receptor CCR-5 -
PT useful to diagnose, prevent and/or treat inflammatory disorders,
PT autoimmune disease and viral infection
XX
XX
PS Claim 7; Fig 1d-e; 94pp; English.
XX
CC The present sequence is an inactive human CC (Cys-Cys)
CC chemokine receptor 5 (CCR5), which lacks the last 3 transmembrane
CC regions and the regions involved in G protein-coupling. CCR5 or
CC its cDNA can be used to diagnose, treat and/or prevent inflammatory
CC diseases, e.g. rheumatoid arthritis, glomerulonephritis, asthma,
CC idiopathic pulmonary fibrosis and psoriasis, viral infections,
CC especially human immunodeficiency virus type 1 or type 2 (HIV-1 or
CC HIV-2) infection, cancer, atherosclerosis and autoimmune disorders.
CC Subjects that express the inactive receptor have a predisposition,
CC or resistance to HIV-1 and/or HIV-2.
XX
SQ Sequence 215 AA;

Query Match 100.0%; Score 1122; DB 18; Length 215;
Best Local Similarity 100.0%; Pred. No. 1.4e-125;
Matches 215; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDYQVSSPIYDINTYTSEPCQKINVKQIAARLLPPLYSLVFIFGVGNMLVILLINCKR 60
DQ |||||||
DB 1 MDYQVSSPIYDINTYTSEPCQKINVKQIAARLLPPLYSLVFIFGVGNMLVILLINCKR 60
QY 61 LKSMTDIYLLNLAIISDLFFLLTPVFWAHYAAQWDFGNTMCOLLTGLYFIFGFSGIFFI 120
DQ |||||||
DB 61 LKSMTDIYLLNLAIISDLFFLLTPVFWAHYAAQWDFGNTMCOLLTGLYFIFGFSGIFFI 120
QY 121 LLTIDRYLAVHVAFAKARTVFGVTVSVITWVAVFASLPGLIFTRSQEGLHYTCSS 180
DQ |||||||
DB 121 LLTIDRYLAVHVAFAKARTVFGVTVSVITWVAVFASLPGLIFTRSQEGLHYTCSS 180
QY 181 HPFYIKDHLGAGPAAACHGHLNLPKNSASVSK 215
DQ |||||||
DB 181 HPFYIKDHLGAGPAAACHGHLNLPKNSASVSK 215

RESULT 2
AAW88238
ID AAW88238 standard; Protein; 215 AA.
XX
AC AAW88238;
XX
DT 15-MAR-1999 (first entry)
XX
DE HIV-1 co-receptor CCR5 variant CCR5-delta32.
XX
KW HIV-1; CCR5; CCR5-delta32; co-receptor; infection; diagnosis; AIDS;
KW gene therapy; human.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Domain 32..56
FT /note= "transmembrane domain 1"
FT Domain 67..87
FT /note= "transmembrane domain 2"
FT Domain 103..124
FT /note= "transmembrane domain 3"
FT Domain 142..167
FT /note= "transmembrane domain 4"
XX
PN W09854317-A1.
XX
PD 03-DEC-1998.
XX
PF 29-MAY-1998; 98WO-EP03437.
XX

```

```

PR 30-MAY-1997; 97US-0048057.
XX
PA (MOND-) FOND MONDIALE RECH & PREVENTION SIDA.
XX
PI Arenzana Siesdedos F, Beretta A, Braun J, Quillent C;
XX
DR WPI; 1999-059835/05.
DR N-PSDB; AAW84159.
XX
XX New CCR5 variant protein of the HIV-1 co-receptor - useful in
PT developing resistance of CCR5-expressing cells to HIV-1 infection
PT
XX Disclosure; Page 38-39; 55pp; English.
XX
CC This is the amino acid sequence of a CCR5 variant protein,
CC designated CCR5-delta32, that includes the first 4 transmembrane
CC domains of wild-type CCR5 (see AAW88232), but lacks transmembrane
CC domains 5-7. CCR5 serves as a co-receptor for infection by
CC macrophage-tropic (M-tropic) strains of HIV-1. Individuals
CC homozygous for the CCR5-delta32 mutation are resistant to HIV-1
CC infection, but heterozygous individuals are susceptible. The
CC invention additionally relates to the identification of variant
CC CCR5s303 (see AAW88231), which lacks transmembrane domains 3-7 of
CC CCR5. The detection of CCR5 variants may be used to identify
CC individuals at lower risk of infection relative to the general
CC population who, if infected, may exhibit slower progression to
CC AIDS. Probes and primers (see AAW84127-36) are provided for use in
CC diagnostic methods for detecting the presence of such variants. A
CC method is provided for inhibiting HIV-1 infection of a cell
CC expressing the CCR5 receptor. This involves introducing a nucleic
CC acid encoding a CCR5 variant into the cell, thereby reducing the
CC number of functional CCR5 molecules present on the cell surface.
XX
SQ Sequence 215 AA;

Query Match 100.0%; Score 1122; DB 20; Length 215;
Best Local Similarity 100.0%; Pred. No. 1.4e-125;
Matches 215; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDYQVSSPIYDINTYTSEPCQKINVKQIAARLLPPLYSLVFIFGVGNMLVILLINCKR 60
DQ |||||||
DB 1 MDYQVSSPIYDINTYTSEPCQKINVKQIAARLLPPLYSLVFIFGVGNMLVILLINCKR 60
QY 61 LKSMTDIYLLNLAIISDLFFLLTPVFWAHYAAQWDFGNTMCOLLTGLYFIFGFSGIFFI 120
DQ |||||||
DB 61 LKSMTDIYLLNLAIISDLFFLLTPVFWAHYAAQWDFGNTMCOLLTGLYFIFGFSGIFFI 120
QY 121 LLTIDRYLAVHVAFAKARTVFGVTVSVITWVAVFASLPGLIFTRSQEGLHYTCSS 180
DQ |||||||
DB 121 LLTIDRYLAVHVAFAKARTVFGVTVSVITWVAVFASLPGLIFTRSQEGLHYTCSS 180
QY 181 HPFYIKDHLGAGPAAACHGHLNLPKNSASVSK 215
DQ |||||||
DB 181 HPFYIKDHLGAGPAAACHGHLNLPKNSASVSK 215

RESULT 3
AAW27406
ID AAW27406 standard; Protein; 184 AA.
XX
AC AAW27406;
XX
DT 14-APR-1998 (first entry)
XX
DE Inactive human CCR5.
XX
KW Inactive; human Cys-Cys chemokine receptor-5; CCR5;
KW human immunodeficiency virus; type 1; type 2; HIV-1; HIV-2;
KW predisposition; resistance; diagnosis; treatment; prevention;
KW inflammatory disease; rheumatoid arthritis; glomerulonephritis;
KW asthma; idiopathic pulmonary fibrosis; psoriasis; viral infection;
KW cancer; atherosclerosis; autoimmune disorder.
XX

```

OS Homo sapiens.
 PN W09732019-A2.
 XX PD 04-SEP-1997.
 XX PF 28-FEB-1997; 97WO-BE00023.
 XX PR 06-AUG-1996; 96EP-0870102.
 XX PR 01-MAR-1996; 96EP-0870021.
 XX PA (EURO-) EUROSREEN SA.
 XX PI Libert F, Parmentier M, Samson M, Vassart G;
 XX WPI; 1997-479829/44.
 XX DR N-PSDB; AAT90116.
 XX PT Active and inactive forms of human CC chemokine receptor CCR-5 -
 PT useful to diagnose, prevent and/or treat inflammatory disorders,
 PT autoimmune disease and viral infection
 XX Claim 1; Fig 1a; 94pp; English.
 XX PS The present sequence is an inactive human CC (Cys-Cys)
 CC chemokine receptor 5 (CCR5), which is not a receptor of human
 CC immunodeficiency virus type 1 or type 2 (HIV-1 or HIV-2). CCR5 or
 CC its cDNA can be used to diagnose, treat and/or prevent inflammatory
 CC diseases, e.g. rheumatoid arthritis, glomerulonephritis, asthma,
 CC idiopathic pulmonary fibrosis and psoriasis, viral infections,
 CC especially HIV-1 or HIV-2 infection, cancer, atherosclerosis and
 CC autoimmune disorders. Subjects that express the inactive receptor
 CC have a predisposition, or resistance to HIV-1 and/or HIV-2.
 XX SQ Sequence 184 AA;
 Query Match 85.4%; Score 958; DB 18; Length 184;
 Best Local Similarity 100.0%; Pred. No. 4.4e-106;
 Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MDQVSSPIYDINVTSEPCOKINVKQIAARLLPPLYSLVFIFGFGNMLVILLINCKR 60
 Db 1 MDQVSSPIYDINVTSEPCOKINVKQIAARLLPPLYSLVFIFGFGNMLVILLINCKR 60
 QY 61 LKSMTDIYLLNLALSDFLLTPFWAHYAAQWDFGNTMCGLLTGLYFIFGSGIFII 120
 Db 61 LKSMTDIYLLNLALSDFLLTPFWAHYAAQWDFGNTMCGLLTGLYFIFGSGIFII 120
 QY 121 LITDRYLVAVHAFKARTVFGVTSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
 Db 121 LITDRYLVAVHAFKARTVFGVTSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
 QY 181 HPFY 184
 Db 181 HPFY 184
 RESULT 4
 AAW27407
 ID AAW27407 standard; Protein; 352 AA.
 XX AC AAW27407;
 XX DT 14-APR-1998 (first entry)
 XX DE Human CCR5.
 XX KW Human Cys-Cys chemokine receptor 5; CCR5;
 KW human immunodeficiency virus; type 1; type 2; HIV-1; HIV-2;
 KW diagnosis; treatment; prevention;
 KW inflammatory disease; rheumatoid arthritis; glomerulonephritis;
 KW asthma; idiopathic pulmonary fibrosis; psoriasis; viral infection;
 KW cancer; atherosclerosis; autoimmune disorder.

XX OS Homo sapiens.
 XX PN W09732019-A2.
 XX PD 04-SEP-1997.
 XX PF 28-FEB-1997; 97WO-BE00023.
 XX PR 06-AUG-1996; 96EP-0870102.
 XX PR 01-MAR-1996; 96EP-0870021.
 XX PA (EURO-) EUROSREEN SA.
 XX PI Libert F, Parmentier M, Samson M, Vassart G;
 XX WPI; 1997-479829/44.
 XX DR N-PSDB; AAT90117.
 XX PT Active and inactive forms of human CC chemokine receptor CCR-5 -
 PT useful to diagnose, prevent and/or treat inflammatory disorders,
 PT autoimmune disease and viral infection
 XX Claim 4; Fig 1b-c; 94pp; English.
 XX PS The present sequence is human CC (Cys-Cys) chemokine receptor
 CC 5 (CCR5), which is stimulated by MIP-1 alpha, MIP-1 beta or RANTES
 CC chemokines, but not by monocyte chemoattractant protein 1 (MCP-1),
 CC MCP-2, MCP-3, interleukin-8 (IL-8) or growth related gene product
 CC alpha (GRO alpha) chemokines. Active CCR-5 is also a receptor of
 CC human immunodeficiency virus type 1 or type 2 (HIV-1 or HIV-2).
 CC CCR5 or its cDNA can be used to diagnose, treat and/or prevent
 CC inflammatory diseases, e.g. rheumatoid arthritis,
 CC glomerulonephritis, asthma, idiopathic pulmonary fibrosis and
 CC psoriasis, viral infections, especially HIV-1 or HIV-2 infection,
 CC cancer, atherosclerosis and autoimmune disorders.
 XX SQ Sequence 352 AA;
 Query Match 85.4%; Score 958; DB 18; Length 352;
 Best Local Similarity 100.0%; Pred. No. 1e-105;
 Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MDQVSSPIYDINVTSEPCOKINVKQIAARLLPPLYSLVFIFGFGNMLVILLINCKR 60
 Db 1 MDQVSSPIYDINVTSEPCOKINVKQIAARLLPPLYSLVFIFGFGNMLVILLINCKR 60
 QY 61 LKSMTDIYLLNLALSDFLLTPFWAHYAAQWDFGNTMCGLLTGLYFIFGSGIFII 120
 Db 61 LKSMTDIYLLNLALSDFLLTPFWAHYAAQWDFGNTMCGLLTGLYFIFGSGIFII 120
 QY 121 LITDRYLVAVHAFKARTVFGVTSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
 Db 121 LITDRYLVAVHAFKARTVFGVTSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
 QY 181 HPFY 184
 Db 181 HPFY 184
 RESULT 5
 AAW27123
 ID AAW27123 standard; Protein; 352 AA.
 XX AC AAW27123;
 XX DT 14-DEC-1997 (first entry)
 XX DE Human chemokine receptor 88C.
 XX KW Chemokine receptor 88C; atherosclerosis; rheumatoid arthritis;
 KW tumour; asthma; viral infection; AIDS; inflammation;
 KW autoimmune disease; therapy; diagnosis; leukocyte trafficking;

G protein coupled receptor; ligand; modulator; antibody; human.
Homo sapiens.
Key Location/Qualifiers
Domain 1..32
/label= Extracellular_domain
Domain 56..67
/label= Intracellular_domain
Domain 89..112
/label= Extracellular_domain
Domain 125..145
/label= Intracellular_domain
Domain 166..191
/label= Extracellular_domain
Domain 213..235
/label= Intracellular_domain
Domain 259..280
/label= Extracellular_domain
Domain 301..352
/label= Intracellular_domain
WO9722698-A2.
26-JUN-1997.
20-DEC-1996; 96WO-US20759.
07-JUN-1996; 96US-0661393.
20-DEC-1995; 95US-0575967.
(ICOS-) ICOS CORP.
Gray PW, Raport CJ, Schweickart VL;
WPI: 1997-341689/31.
N-PSDB; AAT85161.
New nucleic acid encoding chemokine receptors 88-2B and 88C - used to modulate leukocyte trafficking, e.g. for treatment of inflammation, tumours, viral infections, autoimmune diseases, etc.
Claim 16; Page 47-48; 65pp; English.
This polypeptide sequence comprises novel human chemokine receptor 88C, a G protein coupled receptor that is involved in leukocyte trafficking. Its amino sequence was deduced from a cDNA clone (AAT85161) isolated from a macrophage library. It shows 62% identity to CCR1. Chemokine receptor 88-2B (see AAW27124) has also been identified. 88C and 88-2B receptors and their polypeptide fragments can be produced in transformed host cells. The receptors, peptides comprising one or more of the extracellular or intracellular domains, and anti-receptor antibodies can be used to modulate receptor activities, particularly ligand and G protein binding, and are potentially potentially useful in the treatment of atherosclerosis, rheumatoid arthritis, tumours, asthma, viral infection, AIDS, inflammatory conditions, pathological immune response, abnormal haematopoietic processes etc.
Sequence 352 AA;
Query Match 85.4%; Score 958; DB 18; Length 352;
Best Local Similarity 100.0%; Pred. No. 1e-105;
Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDYQSSPTDINYYTSEPCQKINVKQIAARLLPPLYSILVFYGVGNMLVILINCKR 60
DB 1 MDYQSSPTDINYYTSEPCQKINVKQIAARLLPPLYSILVFYGVGNMLVILINCKR 60
QY 61 LKSMTDIYLLNLAISDLFLLTPFVFAHAAAQWDFGNTMCQLLTGLYFGFFSIFII 120
DB 61 LKSMTDIYLLNLAISDLFLLTPFVFAHAAAQWDFGNTMCQLLTGLYFGFFSIFII 120

QY 121 LITDRLVAVHAFALKAARTVTFGVVITWVAVFASLPGLIFTRSQEGLHYTCSS 180
DB 121 LITDRLVAVHAFALKAARTVTFGVVITWVAVFASLPGLIFTRSQEGLHYTCSS 180
QY 181 HFPY 184
DB 181 HFPY 184
RESULT 6
AAW23835
ID AAW23835 standard; Protein; 352 AA.
XX
AC AAW23835;
XX
DT 08-JUN-1998 (first entry)
XX
DE Human CC chemokine receptor 5 (CCR5).
XX
KW CC chemokine receptor 5; CCR5; G-protein coupled receptor;
KW human immunodeficiency virus; HIV; CD4; AIDS; therapy;
KW transgenic animal.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Domain 29..55
FT /label= I
FT /note= "transmembrane domain"
FT Region 109..120
FT /note= "extracellular loop-1 (Claim 19)"
FT Domain 104..126
FT /label= III
FT /note= "transmembrane domain"
FT Domain 143..171
FT /label= IV
FT /note= "transmembrane domain"
FT Region 187..210
FT /note= "extracellular loop-2 (Claim 19)"
FT Domain 194..219
FT /label= V
FT /note= "transmembrane domain"
FT Domain 238..258
FT /label= VI
FT /note= "transmembrane domain"
FT Region 261..276
FT /note= "extracellular loop-3 (Claim 19)"
FT Domain 277..300
FT /label= VII
FT /note= "transmembrane domain"
XX
WO9745543-A2.
XX
PD 04-DEC-1997.
XX
PF 28-MAY-1997; 97WO-US09586.
XX
PR 28-MAY-1996; 96US-0018508.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Alkhatib G, Berger EA, Broder CC, Combadiere C;
PI Feng Y, Kennedy PE, Murphy PM;
XX
DR WPI: 1998-032650/03.
DR N-PSDB; AAT76920.
XX
PT CC chemokine receptor 5 polypeptide - used to inhibit membrane
PT fusion between HIV and a target cell
XX
PS Claim 68; Fig 1C; 70pp; English.
XX
CC This protein sequence comprises of a novel human macrophage-selective

CC chemokine receptor that has been designated CCR5. The sequence
 CC was deduced from an isolated cDNA clone (see AAF76920). An Ala127Ileu
 CC variant (see W238340 of CCR5) was also identified. The susceptibility
 CC of human macrophages to HIV infection depends on cell surface
 CC expression of CD4 and CCR5. CCR5 is a member of the 7-transmembrane
 CC superfamily of G-protein coupled cell surface molecules. It plays
 CC an essential role in the membrane fusion step of infection by some
 CC HIV isolates. The establishment of stable, non-human cell lines
 CC and transgenic mammals having cells that coexpress human CD4 and
 CC CCR5 provides valuable tools for research of HIV infection.
 CC Antibodies that bind to CCR5, CCR5 variants, and CCR5-binding
 CC agents capable of blocking membrane fusion between HIV and target
 CC cells represent potential anti-HIV therapeutics for macrophage
 CC tropic strains of HIV.

XX
 SQ Sequence 352 AA;

Query Match 85.4%; Score 958; DB 19; Length 352;
 Best Local Similarity 100.0%; Pred. No. 1e-105;
 Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDYQVSSPIYDINYYTSEPCQKINVKQIAARLLPPLYSLVIFGFGVGNMVLILINCKR 60
 DB 1 MDYQVSSPIYDINYYTSEPCQKINVKQIAARLLPPLYSLVIFGFGVGNMVLILINCKR 60
 QY 61 LKSMTDIYLLNLAIISDLFLFLLTPFWAHYAAQWDFGNTMCGLLTGLYFGFGSIFPII 120
 DB 61 LKSMTDIYLLNLAIISDLFLFLLTPFWAHYAAQWDFGNTMCGLLTGLYFGFGSIFPII 120
 QY 121 LITIDRYLVAVHAFVAFKARTVFGVTSVITWVAVFASLPGLIIFTRSQKGLHYTCSS 180
 DB 121 LITIDRYLVAVHAFVAFKARTVFGVTSVITWVAVFASLPGLIIFTRSQKGLHYTCSS 180
 QY 181 HPFY 184
 DB 181 HPFY 184

RESULT 7
 AAW88232
 ID AAW88232 standard; Protein; 352 AA.

XX
 AC AAW88232;
 DT 15-MAR-1999 (first entry)
 DE HIV-1 co-receptor CCR5.
 XX
 KW HIV-1; CCR5; CCR5m303; co-receptor; infection; diagnosis; AIDS;
 XX gene therapy; human.
 XX Homo sapiens.

PH Key Location/Qualifiers
 FT Domain 32..56
 FT /note= "transmembrane domain 1"
 FT Domain 67..87
 FT /note= "transmembrane domain 2"
 FT Misc-difference 101
 FT /note= "corresponds to TGR (Cys) in wild-type CCR5,
 FT TGA (Stop) in CCR5m303"
 FT Domain 103..124
 FT /note= "transmembrane domain 3"
 FT Domain 142..167
 FT /note= "transmembrane domain 4"
 FT Domain 200..223
 FT /note= "transmembrane domain 5"
 FT Domain 236..260
 FT /note= "transmembrane domain 6"
 FT Domain 275..301
 FT /note= "transmembrane domain 7"

XX
 PN W09854317-AL.

XX
 PD 03-DEC-1998.
 XX
 PF 29-MAY-1998; 98MO-EP03437.
 XX
 PR 30-MAY-1997; 97US-0048057.
 XX
 PA (MOND-) FOND MONDIALE RECH & PREVENTION SIDA.
 XX
 PI Arenzana Siesdedos F, Beretta A, Braun J, Quillent C;
 XX WPI: 1999-059835/05.
 DR N-PSDB; AAV841126.
 XX
 PT New CCR5 variant protein of the HIV-1 co-receptor - useful in
 XX developing resistance of CCR5-expressing cells to HIV-1 infection
 PS Disclosure; Page 34-35; 55pp; English.
 XX
 CC This is the amino acid sequence of wild-type human CCR5, which
 CC serves as a co-receptor for infection by macrophage-tropic
 CC (M-tropic) strains of HIV-1. The invention relates to the
 CC identification of a CCR5 variant (see AAW88231), designated CCR5m303,
 CC comprising the first two transmembrane domains of wild-type CCR5,
 CC but lacking transmembrane domains 3-7. The presence of the
 CC CCR5m303 variant with the wild type CCR5 allele shows a positive
 CC correlation with resistance to infection with M-tropic HIV-1
 CC strains, and may indicate slower progression of the disease. The
 CC detection of CCR5 variants may be used to identify individuals at
 CC lower risk of infection relative to the general population who, if
 CC infected, may exhibit slower progression to AIDS. Probes and
 CC primers (see AAW84127-36) are provided for use in diagnostic methods
 CC for detecting the presence of such variants. A method is provided
 CC for inhibiting HIV-1 infection of a cell expressing the CCR5
 CC receptor. This involves introducing a nucleic acid encoding a CCR5
 CC variant into the cell, thereby reducing the number of functional
 CC CCR5 molecules present on the cell surface.

XX
 SQ Sequence 352 AA;

Query Match 85.4%; Score 958; DB 20; Length 352;
 Best Local Similarity 100.0%; Pred. No. 1e-105;
 Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDYQVSSPIYDINYYTSEPCQKINVKQIAARLLPPLYSLVIFGFGVGNMVLILINCKR 60
 DB 1 MDYQVSSPIYDINYYTSEPCQKINVKQIAARLLPPLYSLVIFGFGVGNMVLILINCKR 60
 QY 61 LKSMTDIYLLNLAIISDLFLFLLTPFWAHYAAQWDFGNTMCGLLTGLYFGFGSIFPII 120
 DB 61 LKSMTDIYLLNLAIISDLFLFLLTPFWAHYAAQWDFGNTMCGLLTGLYFGFGSIFPII 120
 QY 121 LITIDRYLVAVHAFVAFKARTVFGVTSVITWVAVFASLPGLIIFTRSQKGLHYTCSS 180
 DB 121 LITIDRYLVAVHAFVAFKARTVFGVTSVITWVAVFASLPGLIIFTRSQKGLHYTCSS 180
 QY 181 HPFY 184
 DB 181 HPFY 184

RESULT 8
 ABB56342
 ID ABB56342 standard; Protein; 352 AA.
 XX
 AC ABB56342;
 XX
 DT 18-FEB-2002 (first entry)
 XX
 DE Non-endogenous human GPCR protein, SEQ ID NO: 477.
 XX
 KW Human; G protein-coupled receptor; GPCR; non-endogenous; mutant;
 XX constitutively activated GPCR; agonist; disease.

OS Homo sapiens.
OS Synthetic.
XX WO200177172-A2.
XX PD 18-OCT-2001.
XX PF 05-APR-2001; 2001WO-US11098.
XX PR 07-APR-2000; 2000US-195747P.
XX PA (AREN-) ARENA PHARM INC.
XX PI Lehmann-Bruinsma K, Liaw CW, Lin I;
XX DR WPI; 2001-648759/74.
XX DR N-FSD8; AB197978.
XX PT Identifying agonists of G protein-coupled receptors (GPCRs) for use in
XX PT disease treatment, comprises contacting candidate compounds with
XX PT versions of GPCRs -
XX PS Claim 1; Page 277-278; 394pp; English.
XX CC The invention relates to G protein-coupled receptors (GPCRs) for which
XX CC the endogenous ligand has been identified. Non-endogenous
XX CC constitutively activated versions of known GPCRs are used in the
XX CC invention for the direct identification of candidate compounds as
XX CC receptor agonists, inverse agonists or partial agonists. Such
XX CC agonists are useful as therapeutic agents for diseases or disorders
XX CC associated with GPCRs. The present sequence is a non-endogenous
XX CC version of a known human GPCR.
XX SQ Sequence 352 AA;
Query Match 85.4%; Score 958; DB 22; Length 352;
Best Local Similarity 100.0%; Pred. No. 1e-105;
Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDYQVSSPIYDINTYTSEPCQKINVKQIAARLLPPLYSIVIFGVGNMVLILLINCKR 60
Db 1 MDYQVSSPIYDINTYTSEPCQKINVKQIAARLLPPLYSIVIFGVGNMVLILLINCKR 60
QY 61 LKSMTDIYLLNLAIISDLFFLLTPFWAHYAAAQWDFGNTMQLTGLYFGFSGIFFII 120
Db 61 LKSMTDIYLLNLAIISDLFFLLTPFWAHYAAAQWDFGNTMQLTGLYFGFSGIFFII 120
QY 121 LLTIDRYLAVVHAFKARTVTGVVTSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
Db 121 LLTIDRYLAVVHAFKARTVTGVVTSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
QY 181 HFPY 184
Db 181 HFPY 184
RESULT 9
AAG80111
ID AAG80111 standard; Protein; 352 AA.
XX AC AAG80111;
XX DT 17-JAN-2002 (first entry)
XX DE Human CCR5 protein.
KW Chemokine; tumour diagnosis; colorectal; prostatic; organ rejection;
KW inflammation; autoimmune disease; metastasis; bronchial asthma; lupus;
KW chronic bowel inflammation; rheumatoid arthritis; cytostatic;
KW antiinflammatory; antiasthmatic; immunosuppressive; dermatological;
KW antirheumatic; antiarthritic.

OS Homo sapiens.
XX PN WO200172830-A2.
XX PD 04-OCT-2001.
XX PF 02-APR-2001; 2001WO-EP03708.
XX PR 31-MAR-2000; 2000DE-1016013.
XX PA (IPFP-) IPF PHARM GMBH.
XX PA (FORS/) FORSMANN U.
XX PI Forssmann W, Adermann K, Heitland A, Spodsborg N;
XX DR WPI; 2001-626256/72.
XX PT Diagnostic agent containing two or more receptor-specific ligands,
XX PT useful for detecting tumors, inflammation etc., also therapeutic use of
XX PT ligand inhibitors -
XX PS Disclosure; Page 10; 26pp; German.
XX CC This invention describes a novel diagnostic agent (A) comprising at least
XX CC two different ligands (I) for receptors (II) that are implicated in
XX CC disease. (A) are used for the diagnosis of tumors (especially colorectal
XX CC or prostatic), organ rejection, inflammation and autoimmune diseases.
XX CC Also inhibitors of (I) are used therapeutically against tumors (and their
XX CC metastases), inflammation (particularly bronchial asthma or chronic bowel
XX CC inflammation), or autoimmune diseases (rheumatoid arthritis or lupus),
XX CC where the (cardio)vascular, lymphatic, respiratory, nervous, digestive,
XX CC endocrine, motor or urogenital systems or skin are affected, and bone
XX CC marrow diseases. The products of the invention are chemokine derivatives
XX CC which have cytostatic, antiinflammatory, antiasthmatic,
XX CC immunosuppressive, dermatological, antirheumatic, antiarthritic.
XX CC Chemokines act on specific tumor and inflammatory cells through a
XX CC constellation of chemokine receptors (CR), which control migration and
XX CC proliferation of these cells. AAG8045-AAG80128 represent human chemokine
XX CC fragments used to illustrate the method of the invention.
XX SQ Sequence 352 AA;
Query Match 85.4%; Score 958; DB 22; Length 352;
Best Local Similarity 100.0%; Pred. No. 1e-105;
Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDYQVSSPIYDINTYTSEPCQKINVKQIAARLLPPLYSIVIFGVGNMVLILLINCKR 60
Db 1 MDYQVSSPIYDINTYTSEPCQKINVKQIAARLLPPLYSIVIFGVGNMVLILLINCKR 60
QY 61 LKSMTDIYLLNLAIISDLFFLLTPFWAHYAAAQWDFGNTMQLTGLYFGFSGIFFII 120
Db 61 LKSMTDIYLLNLAIISDLFFLLTPFWAHYAAAQWDFGNTMQLTGLYFGFSGIFFII 120
QY 121 LLTIDRYLAVVHAFKARTVTGVVTSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
Db 121 LLTIDRYLAVVHAFKARTVTGVVTSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
QY 181 HFPY 184
Db 181 HFPY 184
RESULT 10
AAB82948
ID AAB82948 standard; Protein; 352 AA.
XX AC AAB82948;
XX DT 21-DEC-2001 (first entry)
XX DE Human HIV-1 co-receptor CCR5.
XX

SQ Sequence 352 AA;

Query Match 85.4%; Score 958; DB 22; Length 352;
 Best Local Similarity 100.0%; Pred. No. 1e-105;
 Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDYQVSSPIYDINTYTSEPCQKINVKQIAARLLPPLYSILVFYFGVGNMLVILLINCKR 60
 DB 1 MDYQVSSPIYDINTYTSEPCQKINVKQIAARLLPPLYSILVFYFGVGNMLVILLINCKR 60
 QY 61 LKSMTDIYLLNLAIISDLFFLLTPFPWAHYAAQWDFGNTMCQLLTGLYFGFSGIFPII 120
 DB 61 LKSMTDIYLLNLAIISDLFFLLTPFPWAHYAAQWDFGNTMCQLLTGLYFGFSGIFPII 120
 QY 121 LLTIDRYLAVHVAHFALKARTVTVGVTSVITWVAVFASLPGLIIFRSQKGLHYTCSS 180
 DB 121 LLTIDRYLAVHVAHFALKARTVTVGVTSVITWVAVFASLPGLIIFRSQKGLHYTCSS 180
 QY 181 HFPY 184
 DB 181 HFPY 184

RESULT 12
 AAE07048
 ID AAE07048 standard; Protein; 352 AA.
 AC AAE07048;
 XX
 DT 16-OCT-2001 (first entry)
 DE Human G-protein chemokine receptor (CCR5) HDGNR10 protein #2.
 KW Human; G-protein chemokine receptor; CCR5; HDGNR10; inflammation; HIV;
 KW human immunodeficiency virus; antimicrobial; vasodilator; vulnerability;
 KW cystic; immunosuppressive; neutrotic; neuroprotective; gene therapy;
 KW neurodegenerative disorder; Kaposi's sarcoma; autoimmune disease;
 KW rheumatoid arthritis; cancer; breast; ovary; adrenal gland; bone marrow;
 KW gastrointestinal tract; lung; liver; immune disorder; Addison's disease;
 KW haemolytic anaemia; autoimmune thyroiditis; diabetes mellitus; allergy;
 KW multiple sclerosis; ulcerative colitis; Crohn's disease; wound healing;
 KW cardiovascular disorder; myocardial ischaemia.
 OS Homo sapiens.
 XX
 PN WO200158916-A2.
 XX
 PD 16-AUG-2001.
 XX
 PF 09-FEB-2001; 2001WO-US04153.
 XX
 PR 09-FEB-2000; 2000US-0181258.
 PR 09-MAR-2000; 2000US-0187999.
 PR 22-SEP-2000; 2000US-0234336.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Roschke V, Li Y, Ruben SM;
 XX
 DR WPI: 2001-488966/53.
 DR N-PSDB; AAD13299.
 XX
 XX Isolated nucleic acid encoding a human G-protein chemokine receptor
 (CCR5) HDGNR10 polypeptide, useful for preventing or treating
 autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative
 disorders and neurodegenerative disorders -
 XX
 PS Example 40; Page 504-505; 518pp; English.
 XX
 CC The invention relates to human G-protein chemokine receptor (CCR5)
 CC HDGNR10 polypeptides and polynucleotides. CCR5 HDGNR10 antibodies
 CC are useful for treating, preventing or ameliorating a disease or disorder
 CC associated with inflammation, defective or aberrant chemotaxis of immune

CC cells, HIV infection (such as Pneumocystis carinii pneumonia or Kaposi's
 CC sarcoma) or defective or aberrant T-cell antigen presenting cell
 CC interaction. The disease or disorder may also be an infectious disease
 CC (e.g. a viral infection such as an early stage HIV infection, a
 CC cytomegalovirus infection, or a poxvirus infection), an autoimmune
 CC disease (e.g. rheumatoid arthritis) or a neurodegenerative disorder. The
 CC disease or disorder may be associated with aberrant CCR5 expression, lack
 CC of CCR5 function, aberrant CCR5 ligand expression, or lack of CCR5 ligand
 CC function. CCR5 HDGNR10 protein is used as a food additive or preservative
 CC to increase or decrease storage capabilities. CCR5 HDGNR10 DNA are useful
 CC for chromosome identification and in gene therapy. CCR5 HDGNR10 DNA,
 CC protein, antibodies, agonists and antagonists are also useful in the
 CC diagnosis, treatment and prevention of cancer (breast, ovary, adrenal
 CC gland, bone, bone marrow, gastrointestinal tract, liver, lung,
 CC urogenital); immune disorders (Addison's disease, allergies, autoimmune
 CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
 CC colitis); cardiovascular disorders (myocardial ischaemia) and
 CC wound healing. The present sequence is human CCR5 HDGNR10 protein.

SQ Sequence 352 AA;

Query Match 85.4%; Score 958; DB 22; Length 352;
 Best Local Similarity 100.0%; Pred. No. 1e-105;
 Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDYQVSSPIYDINTYTSEPCQKINVKQIAARLLPPLYSILVFYFGVGNMLVILLINCKR 60
 DB 1 MDYQVSSPIYDINTYTSEPCQKINVKQIAARLLPPLYSILVFYFGVGNMLVILLINCKR 60
 QY 61 LKSMTDIYLLNLAIISDLFFLLTPFPWAHYAAQWDFGNTMCQLLTGLYFGFSGIFPII 120
 DB 61 LKSMTDIYLLNLAIISDLFFLLTPFPWAHYAAQWDFGNTMCQLLTGLYFGFSGIFPII 120
 QY 121 LLTIDRYLAVHVAHFALKARTVTVGVTSVITWVAVFASLPGLIIFRSQKGLHYTCSS 180
 DB 121 LLTIDRYLAVHVAHFALKARTVTVGVTSVITWVAVFASLPGLIIFRSQKGLHYTCSS 180
 QY 181 HFPY 184
 DB 181 HFPY 184

RESULT 13
 AAB83354
 ID AAB83354 standard; Protein; 352 AA.
 XX
 AC AAB83354;
 XX
 DT 09-OCT-2001 (first entry)
 DE Human CCR5 protein sequence.
 XX
 XX Chemotactic chemokine receptor 5; gp120; CD4; therapy; HIV; CCR5;
 KW human immunodeficiency virus; anti-inflammatory disease; human.
 XX
 OS Homo sapiens.
 XX
 PN EP1118858-A2.
 XX
 PD 25-JUL-2001.
 XX
 XX 03-JAN-2001; 2001EP-0300020.
 PF
 XX 12-JAN-2000; 2000GB-0000659.
 PR 12-JAN-2000; 2000GB-0000661.
 PR 12-JAN-2000; 2000GB-0000663.
 XX
 PA (PFIZ) PFIZER LTD.
 PA (PFIZ) PFIZER INC.
 XX
 PI Dobbs S, Perros M, Rickett GA;
 XX

DR WPI: 2001-477085/52.
 DR N-PSDB; AAF87099.
 XX
 PT Determining if an agent can modulate CCR5-gp120 interaction, comprises
 PT incubating the agent with CCR5 and gp120 and determining if the agent
 PT modulates the interaction -
 XX
 PS Claim 1; Page 110; 113pp; English.
 XX
 CC This sequence represents the human CCR5 protein sequence.
 CC The invention relates to a method for determining whether an agent is
 CC capable of modulating the interaction of chemotactic chemokine receptor 5
 CC (CCR5) with gp120, comprising incubating the agent with CCR5 and gp120
 CC and determining whether the agent modulates the interaction, where gp120
 CC is associated with CD4, and where the interaction is a low affinity
 CC binding. The method is used to identify an agent capable of modulating
 CC the interaction of CCR5 with gp120. An agent identified by the method
 CC is used to prepare a pharmaceutical composition for the treatment of a
 CC disease or condition associated with CCR5 and gp120 interaction, to treat
 CC a subject with a disease or condition associated with CCR5 and gp120
 CC interaction, and for preparing a pharmaceutical for treating human
 CC immunodeficiency virus (HIV). It can also be used to treat anti-inflammatory
 CC diseases. The method is commercially useful, amenable to high throughput
 CC screening, and detects interaction of gp120 with cells expressing only
 CC CCR5.
 XX
 SQ Sequence 352 AA;

Query Match 85.4%; Score 958; DB 22; Length 352;
 Best Local Similarity 100.0%; Pred. No. 1e-105;
 Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MDYQVSSPIYDINVTSEPCQKINVKQIAARLLPPLYSILVFIFGVGNMVLILLINCKR 60
 DB 1 MDYQVSSPIYDINVTSEPCQKINVKQIAARLLPPLYSILVFIFGVGNMVLILLINCKR 60
 QY 61 LKSMTDIYLLNLALISDLFFLLTPFWAHYAAQWDFGNTMQLLTGLYFIFGFSGIFPII 120
 DB 61 LKSMTDIYLLNLALISDLFFLLTPFWAHYAAQWDFGNTMQLLTGLYFIFGFSGIFPII 120
 QY 121 LTTIDRYLVAVHAFALKARTVTFGVTSVTWVAVFASLPGLIIFTRSQEGLHYTCSS 180
 DB 121 LTTIDRYLVAVHAFALKARTVTFGVTSVTWVAVFASLPGLIIFTRSQEGLHYTCSS 180
 QY 181 HPFY 184
 DB 181 HPFY 184

RESULT 14
 AAE04321
 ID AAE04321 standard; Protein; 352 AA.
 XX
 AC AAE04321;
 XX
 XX 04-SEP-2001 (first entry)
 XX
 DE Human chemokine receptor (CCR), CC-CR-5 related protein #2.
 XX
 KW Human; transformed mammalian cell; CD4; reporter gene; translocation;
 KW human immunodeficiency virus; HIV; long terminal repeat; LTR; therapy;
 KW chemokine receptor; CCR; cellular dysfunction; HIV infection; cofactor;
 KW CC-CR-5; envelope glycoprotein; anti-HIV.
 XX
 OS Homo sapiens.
 XX
 XX US6258527-B1.
 XX
 PD 10-JUL-2001.
 XX
 XX 21-MAY-1997; 97US-0861105.
 XX
 XX 20-MAY-1996; 96US-0017157.
 PR

PR 19-JUN-1996; 96US-0020043.
 PR 19-MAY-1997; 97US-0858660.
 XX
 XX (AARO-) AARON DIAMOND AIDS RES CENT.
 PA (UNY) UNIV NEW YORK STATE.
 XX
 PI Littman DR, Deng H, Ellmeier W, Landau NR, Liu R;
 XX
 XX WPI; 2001-417127/44.
 DR N-PSDB; AAD08577.
 XX
 PT Transformed mammalian cell (I) that contains a CD4 gene, reporter gene
 PT and HIV LTR for identification of drugs and antibodies for treatment of
 PT HIV -
 XX
 PS Disclosure; Column 47-50; 37pp; English.
 XX
 CC The present invention relates to a transformed mammalian cell that
 CC contains a gene encoding CD4, a construct encoding a reporter gene under
 CC the regulation of an human immuno deficiency virus (HIV) long terminal
 CC repeat (LTR) and that has been transduced with a vector encoding a
 CC human chemokine receptor (CCR) where the CD4 and the CCR are present on
 CC the cell surface of transformed mammalian cell. The invention is useful
 CC for identifying drugs or antibodies that interfere with the
 CC translocation of HIV into transformed mammalian cell or for identifying
 CC a human chemokine receptor that facilitates the infection of a
 CC particular HIV strain into the transformed mammalian cell. Compounds
 CC identified can be used to treat cellular dysfunction and to prevent or
 CC combat HIV infection. The present sequence is a human chemokine receptor
 CC (CCR), CC-CR-5 related protein. CC-CR-5 is the principal cofactor for
 CC entry mediated by the envelope glycoproteins of primary macrophage-tropic
 CC strains of HIV-1.
 XX
 SQ Sequence 352 AA;

Query Match 85.4%; Score 958; DB 22; Length 352;
 Best Local Similarity 100.0%; Pred. No. 1e-105;
 Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MDYQVSSPIYDINVTSEPCQKINVKQIAARLLPPLYSILVFIFGVGNMVLILLINCKR 60
 DB 1 MDYQVSSPIYDINVTSEPCQKINVKQIAARLLPPLYSILVFIFGVGNMVLILLINCKR 60
 QY 61 LKSMTDIYLLNLALISDLFFLLTPFWAHYAAQWDFGNTMQLLTGLYFIFGFSGIFPII 120
 DB 61 LKSMTDIYLLNLALISDLFFLLTPFWAHYAAQWDFGNTMQLLTGLYFIFGFSGIFPII 120
 QY 121 LTTIDRYLVAVHAFALKARTVTFGVTSVTWVAVFASLPGLIIFTRSQEGLHYTCSS 180
 DB 121 LTTIDRYLVAVHAFALKARTVTFGVTSVTWVAVFASLPGLIIFTRSQEGLHYTCSS 180
 QY 181 HPFY 184
 DB 181 HPFY 184

RESULT 15
 AAB46858
 ID AAB46858 standard; Protein; 352 AA.
 XX
 AC AAB46858;
 XX
 XX 16-AUG-2001 (updated)
 DT 02-AUG-2001 (updated)
 DT 04-MAY-2001 (first entry)
 XX
 DE Human HDGHR10 protein.
 XX
 KW HDGHR10; human; G-protein chemokine receptor; antiinflammatory;
 KW immunomodulatory; anticoagulant; antiallergic; immunosuppressive;
 KW cytostatic; antiparasitic; antipneumatic; antirheumatic; antiarthritic;
 KW vasotropic; gene therapy; haematopoiesis; wound healing; coagulation;
 KW angiogenesis; solid tumour; infection; leukemia; growth factor activity;

KW T-cell mediated autoimmune disease; psoriasis; allergy; atherogenesis;
KW anaphylaxis; malignancy; inflammation; histamine; IgE; silicosis; shock;
KW immunoglobulin E-mediated allergic reaction; rheumatoid arthritis;
KW prostaglandin-independent fever; bone marrow failure; sarcoidosis;
KW hyper-eosinophilic syndrome; vulvarity.

OS Homo sapiens.

XX US2001000241-A1.

XX 12-APR-2001.

XX 29-NOV-2000; 2000US-0725285.

XX 06-JUN-1995; 95US-0466343.

PR 18-NOV-1998; 98US-0195662.

XX 25-JUN-1999; 99US-0339912.

XX (LIY/) LI Y.

PA (RUBE/) RUBEN S M.

XX LI Y, Ruben SM;

XX WPI; 2001-226317/23.

DR N-PSDB; AAF26390.

XX New human G-protein chemokine receptor polypeptides and

PT polynucleotides, useful for identifying (ant)agonists to the G-protein

PT chemokine receptor -

XX Claim 1a; Page 15; 22pp; English.

XX This invention describes a novel receptor polypeptide (I) selected from
CC (i) a fully defined 329 amino acid sequence (ii) fully disclosed in the
CC specification; and (iii) a polypeptide encoded by the cDNA contained in a
CC plasmid, and fragments, analogs and derivatives of the polypeptide. The
CC products of the invention have antiinflammatory, immunomodulatory,
CC anticoagulant, antiallergic, immunosuppressive, vulvarity, cytostatic,
CC antiparasitic, antipsoriatic, antirheumatic, antiarthritic and vasotropic
CC activity and can be used for gene therapy. The G-protein chemokine
CC receptors, HGNR10, (I) are useful for screening for compounds which
CC activate or inhibit activation of (I). The products of the invention can
CC also be used for stimulating haematopoiesis, wound healing, coagulation,
CC angiogenesis, treating solid tumours, chronic infections, leukemia,
CC T-cell mediated autoimmune diseases, parasitic infections, psoriasis, and
CC stimulating growth factor activity. HGNR10 is useful for treating
CC allergy, atherogenesis, anaphylaxis, malignancy, chronic and acute
CC inflammation, histamine and immunoglobulin E (IgE)-mediated allergic
CC reactions, prostaglandin-independent fever, bone marrow failure,
CC silicosis, sarcoidosis, rheumatoid arthritis, shock and
CC hyper-eosinophilic syndrome.

CC (N.B. This record was resubmitted to correct errors in the keyword
CC formatting).

XX SQ Sequence 352 AA;

Query Match 85.4%; Score 958; DB 22; Length 352;

Best Local Similarity 100.0%; Pred. No. 1e-105;

Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MDYQVSPFIYDINTYSEPCQKINVKGIAARLLPPLYSLVIFGFVGNMVLILINCKR 60

Db 1 MDYQVSPFIYDINTYSEPCQKINVKGIAARLLPPLYSLVIFGFVGNMVLILINCKR 60

OY 61 LKSMTDIYLLNLAISDLFFLTVPFWAHYAAQWDFGNTMCQLLTGLYFIFGSGIFFII 120

Db 61 LKSMTDIYLLNLAISDLFFLTVPFWAHYAAQWDFGNTMCQLLTGLYFIFGSGIFFII 120

OY 121 LITDRYLAVVHAFVAVFVGVVTVITWVAVFASLPGIIFTRSQEGLHYTCSS 180

Db 121 LITDRYLAVVHAFVAVFVGVVTVITWVAVFASLPGIIFTRSQEGLHYTCSS 180

OY 181 HPPY 184

Db 181 HPPY 184

Search completed: June 3, 2003, 19:22:34
Job time : 56.061 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 3, 2003, 18:49:55 ; Search time 4.22266 Seconds
(without alignments)
4506.655 Million cell updates/sec

Title: US-09-938-703-3_COPY_792_885

Perfect score: 164

Sequence: 1 ATTAAAGATGATCATCTGG.....CTCTGCTGGTGTGCAAT 94

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 383519 seqs, 101223694 residues

Total number of hits satisfying chosen parameters: 767038

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_n2p.model -DEV=xlp
-O=/cgn2_1/USPTO_pool/US0938703/runat_03062003_161410_23015/app_query.fasta_1.1870
-DB=Published_Applications_AA -QMT=frastan -SUFFIX=rapb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOALIGN=200 -THR_SCORE=ptc -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US0938703@cgn.1.1.24 @runat_03062003_161410_23015
-NCPU=6 -ICPU=3 -NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPEXT=0.5
-Fgapop=6 -Fgapext=7 -Ygapop=10 -Ygapext=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications AA:*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|------------------------------------|
| 1 | 164 | 100.0 | 34 | 10 | US-09-938-719-13 Sequence 13, Appl |
| 2 | 164 | 100.0 | 34 | 10 | US-09-939-226-13 Sequence 13, Appl |
| 3 | 164 | 100.0 | 34 | 10 | US-09-938-703-13 Sequence 13, Appl |
| 4 | 164 | 100.0 | 215 | 10 | US-09-938-719-6 Sequence 6, Appl |

| | | | | | | |
|----|-----|-------|-----|----|-------------------|-------------------|
| 5 | 164 | 100.0 | 215 | 10 | US-09-939-226-6 | Sequence 6, Appl |
| 6 | 164 | 100.0 | 215 | 10 | US-09-938-703-6 | Sequence 6, Appl |
| 7 | 152 | 92.7 | 54 | 10 | US-09-938-719-11 | Sequence 11, Appl |
| 8 | 152 | 92.7 | 54 | 10 | US-09-939-226-11 | Sequence 11, Appl |
| 9 | 152 | 92.7 | 54 | 10 | US-09-938-703-11 | Sequence 11, Appl |
| 10 | 152 | 92.7 | 352 | 9 | US-10-232-686-2 | Sequence 2, Appl |
| 11 | 152 | 92.7 | 352 | 9 | US-10-086-814-1 | Sequence 1, Appl |
| 12 | 152 | 92.7 | 352 | 9 | US-09-734-221A-14 | Sequence 14, Appl |
| 13 | 152 | 92.7 | 352 | 10 | US-09-725-285-2 | Sequence 2, Appl |
| 14 | 152 | 92.7 | 352 | 10 | US-09-759-841-2 | Sequence 2, Appl |
| 15 | 152 | 92.7 | 352 | 10 | US-09-779-879A-2 | Sequence 2, Appl |
| 16 | 152 | 92.7 | 352 | 10 | US-09-779-879A-22 | Sequence 22, Appl |
| 17 | 152 | 92.7 | 352 | 10 | US-09-779-880A-2 | Sequence 2, Appl |
| 18 | 152 | 92.7 | 352 | 10 | US-09-779-880A-22 | Sequence 22, Appl |
| 19 | 152 | 92.7 | 352 | 10 | US-09-813-653-15 | Sequence 15, Appl |
| 20 | 152 | 92.7 | 352 | 10 | US-09-813-653-17 | Sequence 17, Appl |
| 21 | 152 | 92.7 | 352 | 10 | US-09-796-202-1 | Sequence 1, Appl |
| 22 | 152 | 92.7 | 352 | 10 | US-09-195-662A-2 | Sequence 2, Appl |
| 23 | 152 | 92.7 | 352 | 10 | US-09-339-912A-2 | Sequence 2, Appl |
| 24 | 152 | 92.7 | 352 | 10 | US-09-938-719-5 | Sequence 5, Appl |
| 25 | 152 | 92.7 | 352 | 10 | US-09-939-226-5 | Sequence 5, Appl |
| 26 | 152 | 92.7 | 352 | 10 | US-09-938-703-5 | Sequence 5, Appl |
| 27 | 152 | 92.7 | 352 | 10 | US-09-502-783A-2 | Sequence 2, Appl |
| 28 | 152 | 92.7 | 352 | 12 | US-10-106-623-2 | Sequence 2, Appl |
| 29 | 149 | 90.9 | 352 | 12 | US-10-106-623-20 | Sequence 20, Appl |
| 30 | 134 | 81.7 | 329 | 10 | US-09-725-285-9 | Sequence 9, Appl |
| 31 | 134 | 81.7 | 329 | 10 | US-09-185-662A-9 | Sequence 9, Appl |
| 32 | 134 | 81.7 | 329 | 10 | US-09-339-912A-9 | Sequence 9, Appl |
| 33 | 134 | 81.7 | 329 | 10 | US-09-502-783A-9 | Sequence 9, Appl |
| 34 | 134 | 81.7 | 344 | 9 | US-10-232-686-9 | Sequence 9, Appl |
| 35 | 134 | 81.7 | 344 | 10 | US-09-779-879A-9 | Sequence 9, Appl |
| 36 | 134 | 81.7 | 344 | 10 | US-09-779-880A-9 | Sequence 9, Appl |
| 37 | 134 | 81.7 | 347 | 10 | US-09-104-792-3 | Sequence 3, Appl |
| 38 | 134 | 81.7 | 360 | 10 | US-09-131-827A-20 | Sequence 20, Appl |
| 39 | 134 | 81.7 | 360 | 10 | US-09-131-827A-2 | Sequence 2, Appl |
| 40 | 134 | 81.7 | 360 | 10 | US-09-938-719-7 | Sequence 7, Appl |
| 41 | 134 | 81.7 | 360 | 10 | US-09-939-226-7 | Sequence 7, Appl |
| 42 | 134 | 81.7 | 360 | 10 | US-09-938-703-7 | Sequence 7, Appl |
| 43 | 109 | 66.5 | 332 | 9 | US-10-001-835-140 | Sequence 140, App |
| 44 | 109 | 66.5 | 355 | 9 | US-09-922-895-1 | Sequence 1, Appl |
| 45 | 109 | 66.5 | 355 | 10 | US-09-931-381A-16 | Sequence 16, Appl |

ALIGNMENTS

RESULT 1
US-09-938-719-13
; Sequence 13, Application US/0938719
; Patent No. US20020106742A1
; GENERAL INFORMATION:
; APPLICANT: SAMSON, MICHEL
; PARMENTIER, MARC
; VASSART, FREDERICK
; LIBERT, GILBERT
; TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR
AND NUCLEIC ACID MOLECULES ENCODING SAID RECEPTOR
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION NUMBER: US/09/938,719
; APPLICATION NUMBER: US/09/938,719
; FILING DATE: 24-Aug-2001
; CLASSIFICATION: <Unknown>

[illegible]

```

RESULT 5
US-09-939-226-6
; Sequence 6, Application US/09939226
; Patent No. US20020110805A1
; GENERAL INFORMATION:
; APPLICANT: SAMSON, MICHEL
; ; PARMENTIER, MARC
; ; VASSART, GILBERT
; ; LIBERT, FREDERICK
; TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR
; AND NUCLEIC ACID MOLECULES ENCODING SAID RECEPTOR
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/939,226
; FILING DATE: 24-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/626,939
; FILING DATE: 2000-07-27
; ATTORNEY/AGENT INFORMATION:
; NAME: Altman, Daniel E
; REGISTRATION NUMBER: 34,115
; REFERENCE/DOCKET NUMBER: <Unknown>
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 215 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-939-226-6

Alignment Scores:
Pred. No.: 5.06e-15 Length: 215
Score: 164.00 Matches: 31
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-09-938-703-3_COPY_792_885 (1-94) x US-09-939-226-6 (1-215)
QY 1 ATTAAGATGATCATCTTGGCGTGTCCGCGTGTCTGTCATGTCATCTGCTACTC 60
Db 185 ILtlyaspSrHisLeuGlyAlaGlyProAlaAlaCysHisGlyHisLeuLeu 204
QY 61 GGCAATCCATAAAACTCTGCTTCGGTGTGCGAAA 93
Db 205 GlyAsnProLysAsnSerAlaSerValSerLys 215

RESULT 6
US-09-938-703-6
; Sequence 6, Application US/09938703
; Patent No. US20020110870A1
; GENERAL INFORMATION:
; APPLICANT: SAMSON, MICHEL
; ; PARMENTIER, MARC
; ; VASSART, GILBERT
; ; LIBERT, FREDERICK
; TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR

```


| | | | |
|---|---|----------|-----------------|
| Alignment Scores: | | 2.33e-13 | Length: 54 |
| Pred. No.: | Score: | 152.00 | Matches: 31 |
| Percent Similarity: | | 100.00% | Conservative: 0 |
| Best Local Similarity: | | 100.00% | Mismatches: 0 |
| Query Match: | | 92.68% | Indels: 0 |
| DB: | | 10 | Gaps: 0 |
| US-09-938-703-3_COPY_792_885 (1-94) x US-09-938-703-11 (1-54) | | | |
| QY | 2 TTAAGATAGTCAATCTGGGGCTGCTCTCCGCTGCTTGTTCATGTCATCTGCTACTCG 61 | | |
| Db | 20 LeuLysleValleleuGlyLeuValleuProleuLeuValMetValleleCysTyrSer 139 | | |
| QY | 62 GGAAFTCTTAAAAACTCTGCTTCGGTGTGGAAT 94 | | |
| Db | 40 GlylleLeuLysThrLeuLeuArgCysArgAsn 50 | | |
| RESULT 10 | | | |
| US-10-232-686-2 | | | |
| ; Sequence 2, Application US/10232686 | | | |
| ; Publication No. US20030023044A1 | | | |
| ; GENERAL INFORMATION: | | | |
| ; APPLICANT: L1, Yi | | | |
| ; APPLICANT: Ruben, Steven M. | | | |
| ; TITLE OF INVENTION: Human G-Protein Chemokine Receptor (CCR5) HDGNR10 | | | |
| ; FILE REFERENCE: 1488.115000N | | | |
| ; CURRENT APPLICATION NUMBER: US/10/232,686 | | | |
| ; CURRENT FILING DATE: 2002-09-03 | | | |
| ; PRIOR APPLICATION NUMBER: 09/339,912 | | | |
| ; PRIOR FILING DATE: 1999-06-25 | | | |
| ; PRIOR APPLICATION NUMBER: 09/195,662 | | | |
| ; PRIOR FILING DATE: 1998-11-18 | | | |
| ; PRIOR APPLICATION NUMBER: 08/466,343 | | | |
| ; PRIOR FILING DATE: 1995-06-06 | | | |
| ; NUMBER OF SEQ ID NOS: 9 | | | |
| ; SOFTWARE: PatentIn version 3.0 | | | |
| ; SEQ ID NO 2 | | | |
| ; LENGTH: 352 | | | |
| ; TYPE: PRT | | | |
| ; ORGANISM: Homo sapiens | | | |
| US-10-232-686-2 | | | |
| Alignment Scores: | | 2.68e-13 | Length: 352 |
| Pred. No.: | Score: | 152.00 | Matches: 31 |
| Percent Similarity: | | 100.00% | Conservative: 0 |
| Best Local Similarity: | | 100.00% | Mismatches: 0 |
| Query Match: | | 92.68% | Indels: 0 |
| DB: | | 9 | Gaps: 0 |
| US-09-938-703-3_COPY_792_885 (1-94) x US-10-232-686-2 (1-352) | | | |
| QY | 2 TTAAGATAGTCAATCTGGGGCTGCTCTCCGCTGCTTGTTCATGTCATCTGCTACTCG 61 | | |
| Db | 196 LeuLysleValleleuGlyLeuValleuProleuLeuValMetValleleCysTyrSer 219 | | |
| QY | 62 GGAAFTCTTAAAAACTCTGCTTCGGTGTGGAAT 94 | | |
| Db | 216 GlylleLeuLysThrLeuLeuArgCysArgAsn 226 | | |
| RESULT 11 | | | |
| US-10-086-814-1 | | | |
| ; Sequence 1, Application US/10086814 | | | |
| ; Publication No. US20030092632A1 | | | |
| ; GENERAL INFORMATION: | | | |
| ; APPLICANT: Dragic, Tatjana | | | |
| ; APPLICANT: Olson, William C. | | | |
| ; TITLE OF INVENTION: SULFATED CCR5 PEPTIDES FOR HIV-1 INFECTION | | | |
| ; FILE REFERENCE: 61010-AB-1 | | | |
| ; CURRENT APPLICATION NUMBER: US/10/086,814 | | | |
| ; CURRENT FILING DATE: 2002-02-28 | | | |

;

TYPE: amino acid

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2003, 19:19:16 ; Search time 6.42683 Seconds
(without alignments)
993.874 Million cell updates/sec

Title: US-09-938-703-6_COPY_185_215
Perfect score: 164
Sequence: 1 IKDHLGAGPAAACHGHLILGNPKNSASVSK 31

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- 1: sp.archaea.*
- 2: sp.bacteria.*
- 3: sp.fungi.*
- 4: sp.human.*
- 5: sp.invertebrate.*
- 6: sp.mammal.*
- 7: sp.mhc.*
- 8: sp.organelle.*
- 9: sp.phage.*
- 10: sp.plant.*
- 11: sp.rodent.*
- 12: sp.virus.*
- 13: sp.vertibrate.*
- 14: sp.unclassified.*
- 15: sp.virus.*
- 16: sp.bacteriap.*
- 17: sp.archaea.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------|
| 1 | 164 | 100.0 | 215 | 4 | 075303 |
| 2 | 115 | 70.1 | 33 | 4 | Q9P1T8 |
| 3 | 55 | 33.5 | 52 | 4 | Q9P1T5 |
| 4 | 55 | 33.5 | 178 | 4 | Q8TEJ8 |
| 5 | 55 | 33.5 | 336 | 16 | Q8VT73 |
| 6 | 54 | 32.9 | 276 | 5 | Q9W5C7 |
| 7 | 54 | 32.9 | 407 | 16 | Q8XNW7 |
| 8 | 54 | 32.9 | 727 | 2 | Q9AF65 |
| 9 | 53 | 32.3 | 312 | 16 | Q9LID3 |
| 10 | 53 | 32.3 | 499 | 16 | Q9A9Q4 |
| 11 | 52.5 | 32.0 | 478 | 10 | Q9LES2 |
| 12 | 52.5 | 32.0 | 478 | 10 | Q94B29 |
| 13 | 52 | 31.7 | 408 | 16 | Q97FK0 |
| 14 | 51 | 31.1 | 141 | 13 | Q9W619 |
| 15 | 51 | 31.1 | 187 | 3 | Q8W2T3 |
| 16 | 50 | 30.5 | 183 | 4 | Q9P098 |

| | | | | | |
|----|------|------|------|----|--------|
| 17 | 49.5 | 30.2 | 196 | 11 | Q8R3C9 |
| 18 | 49.5 | 30.2 | 756 | 11 | Q9CW46 |
| 19 | 49 | 29.9 | 331 | 10 | Q64731 |
| 20 | 49 | 29.9 | 515 | 12 | Q40960 |
| 21 | 49 | 29.9 | 603 | 5 | Q9XVA3 |
| 22 | 48.5 | 29.6 | 500 | 6 | Q9TUJ5 |
| 23 | 48.5 | 29.6 | 698 | 5 | Q9W3F7 |
| 24 | 48 | 29.3 | 242 | 5 | Q16761 |
| 25 | 48 | 29.3 | 325 | 11 | Q9QZP7 |
| 26 | 48 | 29.3 | 370 | 5 | Q95TX4 |
| 27 | 48 | 29.3 | 626 | 11 | Q99KB9 |
| 28 | 48 | 29.3 | 773 | 16 | Q9L1B7 |
| 29 | 48 | 29.3 | 828 | 5 | Q9VKR7 |
| 30 | 48 | 29.3 | 989 | 2 | Q59736 |
| 31 | 48 | 29.3 | 1156 | 3 | Q8SQY1 |
| 32 | 48 | 29.3 | 1223 | 4 | Q8WXS8 |
| 33 | 48 | 29.3 | 1223 | 4 | Q8TE55 |
| 34 | 47.5 | 29.0 | 215 | 13 | Q9P093 |
| 35 | 47.5 | 29.0 | 677 | 10 | Q9FIM7 |
| 36 | 47.5 | 29.0 | 747 | 5 | Q9BIE6 |
| 37 | 47.5 | 29.0 | 971 | 5 | Q9VEC5 |
| 38 | 47 | 28.7 | 159 | 12 | Q997G6 |
| 39 | 47 | 28.7 | 170 | 12 | Q65868 |
| 40 | 47 | 28.7 | 217 | 5 | Q9VL41 |
| 41 | 47 | 28.7 | 246 | 10 | Q94E56 |
| 42 | 47 | 28.7 | 294 | 17 | Q9FD42 |
| 43 | 47 | 28.7 | 302 | 12 | Q36364 |
| 44 | 47 | 28.7 | 307 | 5 | Q9W158 |
| 45 | 47 | 28.7 | 330 | 16 | Q8XRC0 |

ALIGNMENTS

RESULT 1
075303 PRELIMINARY; PRT; 215 AA.
AC 075303;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE CC-chemokine receptor.
GN CCR-5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Tse L., Ehrenberg P.K., Chang G., Michael N.L.;
RT *Genomic Organization and Functional Characterization of the Complete Transcription Unit for the Chemokine Receptor CCR-5, a Major Entry Co-Receptor for HIV-1.*;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF009962; AAC23944.1; -;
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam: PF00001; 7tm_1; 1.
DR PROSITE: PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE: PS00262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 215 AA; 23946 MW; 3C9146C76BA416F7 CRC64;
Query Match 100.0%; Score 164; DB 4; Length 215;
Best Local Similarity 100.0%; Pred. No. 3.9e-16;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 IKDHLGAGPAAACHGHLILGNPKNSASVSK 31
Db 185 IKDHLGAGPAAACHGHLILGNPKNSASVSK 215
RESULT 2
Q9P1T8

```

ID O9P1T8 PRELIMINARY; PRT; 33 AA.
AC Q9P1T8;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
DE Mutant chemokine receptor CCR5 (Fragment).
GN CCR5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99392755; PubMed=10465086;
RA Aikhionbare F.O., Newman C., Womack C., Roth W.W., Stringer H.G. Jr.,
RA Bond V.C.;
RT "Characterization of a third CCR5 amplicon from CCR5-delta32-
RT heterozygous HIV-1-infected individuals.";
RL AIDS 13:1585-1586(1999).
DR EMBL; AF052244; AAF65577.1; -.
KW Receptor.
FT NON_TER 1
SQ SEQUENCE 33 AA; 3477 MW; 664294782EDB3E1 CRC64;

Query Match 70.1%; Score 115; DB 4; Length 33;
Best Local Similarity 100.0%; Pred. No. 9e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKDHLGAGPAAACHGHLILG 21
Db |||||
11 IKDHLGAGPAAACHGHLILG 31

RESULT 3
Q9P1T5
ID O9P1T5 PRELIMINARY; PRT; 52 AA.
AC Q9P1T5;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
DE Mutant chemokine receptor CCR5 (Fragment).
GN CCR5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99392755; PubMed=10465086;
RA Aikhionbare F.O., Newman C., Womack C., Roth W.W., Stringer H.G. Jr.,
RA Bond V.C.;
RT "Characterization of a third CCR5 amplicon from CCR5-delta32-
RT heterozygous HIV-1-infected individuals.";
RL AIDS 13:1585-1586(1999).
DR EMBL; AF056019; AAF65578.1; -.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
KW Receptor.
FT NON_TER 1
FT VARIANT 41 41 P -> S.
SQ SEQUENCE 52 AA; 5962 MW; DAEB2A5A9529C3A9 CRC64;

Query Match 33.5%; Score 55; DB 4; Length 52;
Best Local Similarity 100.0%; Pred. No. 0.93;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 GNPKNASVSK 31
Db |||||
42 GNPKNASVSK 52

RESULT 4
Q8TEJ8

```

```

ID Q8TEJ8 PRELIMINARY; PRT; 178 AA.
AC Q8TEJ8;
DT 01-JUN-2002 (TRENBLrel. 21, Created)
DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE FLJ00197 protein (Fragment).
GN FLJ00197.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=SPLEEN;
RA Jikuya H., Takano J., Nomura N., Kikuno R., Nagase T., Ohara O.;
RT "The nucleotide sequence of a long cDNA clone isolated from human
RT spleen.";
RL Submitted (JAN-2002) to the EMBL/GenBank/DDBI databases.
DR EMBL; AK074125; BAB84951.1; -.
FT NON_TER 1
SQ SEQUENCE 178 AA; 18914 MW; E7BAB999932424CB CRC64;

Query Match 33.5%; Score 55; DB 4; Length 178;
Best Local Similarity 50.0%; Pred. No. 3.3;
Matches 12; Conservative 1; Mismatches 5; Indels 6; Gaps 1;

QY 3 DSHLGAGPAAACHGHLILGNPKNS 26
Db |||||
120 DSHLGAGPAAACHGHLILGNPKNS 137

RESULT 5
Q8YT73
ID Q8YT73 PRELIMINARY; PRT; 336 AA.
AC Q8YT73;
DT 01-MAR-2002 (TRENBLrel. 20, Created)
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE UDP-glucose 4-epimerase.
GN AUL2853.
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21595285; PubMed=11759840;
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
RT cyanobacterium Anabaena sp. strain PCC 7120.";
RL DNA Res. 8:205-213(2001).
DR EMBL; AP003591; BAB74552.1; -.
DR InterPro; IPR001509; Epimerase_Dh.
DR DR InterPro; IPR000205; NAD_binding.
DR Pfam; PF01370; Epimerase; 1.
DR TIGRPFAM; TIGR01179; galE; 1.
KW Complete proteome.
SQ SEQUENCE 336 AA; 36733 MW; 5C4E7F6F6E855276 CRC64;

Query Match 33.5%; Score 55; DB 16; Length 336;
Best Local Similarity 35.5%; Pred. No. 6.3;
Matches 11; Conservative 7; Mismatches 13; Indels 0; Gaps 0;

QY 1 IKDHLGAGPAAACHGHLILGNPKNSVSK 31
Db |||||
31 VYDNCSTGSPQAVLHGELIIGDLKNSCLSQ 61

RESULT 6
Q9W5C7

```

```

ID Q9W5C7 PRELIMINARY; PRT; 276 AA.
AC Q9W5C7;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
DE CG14633 protein.
GN CG14633.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brannon R.C., Rogers Y.-H.C., Blazej R.G., Champagne M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.I.G.,
RA April J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Bencs P.V., Beriman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Rulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
DR EMBL; AE003419; AA45560.1; -.
DR FlyBase; FBgn029537; CG14633.
SQ SEQUENCE 276 AA; 28602 MW; 9F10E7DE62EA3A7 CRC64;

Query Match 32.9%; Score 54; DB 5; Length 276;
Best Local Similarity 42.4%; Pred. No. 7.3;
Matches 14; Conservative 1; Mismatches 2; Indels 16; Gaps 2;

QY 7 GAGPAAAC-----H-GHLLGNP 23
Db 167 GAGPAAACLVCCSPSSHHHGHVGHATGHP 199
|||||
| | | | |

RESULT 7
Q8XNW7 PRELIMINARY; PRT; 407 AA.
AC Q8XNW7;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)

```

```

DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Probable exonuclease.
GN SBCE OR CPE0215.
OS Clostridium perfringens.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Clostridiales; Clostridiaceae; Clostridium.
OX NCBI_TaxID=1502;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=13 / TYPE A;
RX PubMed=11792842;
RA Shimizu T., Ohtani K., Hirakawa H., Ohshima K., Yanashita A.,
RA Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.;
RT "Complete genome sequence of Clostridium perfringens, an anaerobic
RT flesh-eater."
RL Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).
DR EMBL; AP003185; BAB79921.1; -.
DR InterPro; IPR004843; M-peptidase.
DR InterPro; IPR004844; S/T_phosphatase.
DR InterPro; IPR004593; SBCE.
DR Pfam; PF00149; Metallophos; 1.
DR TIGRFAMs; TIGR00619; sbcd; 1.
KW Exonuclease; Complete proteome.
SQ SEQUENCE 407 AA; 46819 MW; 5E71E267DDC8B4A5 CRC64;

Query Match 32.9%; Score 54; DB 16; Length 407;
Best Local Similarity 38.5%; Pred. No. 11;
Matches 10; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

QY 6 LGAGPAAACHGHLILGNPKNSASVSK 31
: | | | | | | | | | | | | | | | | | |
Db 92 VAAGPLAREHGTLINVTGPKSIVEGVE 117

RESULT 8
Q9AF65 PRELIMINARY; PRT; 727 AA.
AC Q9AF65;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Cellulase Cella.
OS Clavibacter michiganensis subsp. sepeodonicus.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Micrococciaceae; Microbacteriaceae; Clavibacter.
OX NCBI_TaxID=31964;
RN [1]
RP SEQUENCE FROM N.A.
RA Laine M., Haapalainen M., Wahlroos T., Kankare K., Missinen R.,
RA Kassuwi S., Metzler M.C.;
RT "The cellulase encoded by the native plasmid of Clavibacter
RT michiganensis subsp. sepeodonicus plays a role in virulence and
RT contains an expansin-like domain."
RL Physiol. Mol. Plant Pathol. 0:0-0(2001).
DR EMBL; AY007311; AAK16222.1; -.
DR HSP; P54583; IECF.
DR InterPro; IPR001919; Bac_cellose-bind.
DR InterPro; IPR001547; GH_5.
DR Pfam; PF00553; CBW_2; 1.
DR Pfam; PF00150; cellulase; 1.
DR PROSITE; PS00659; GLYCOSYL_HYDROL_F5; UNKNOWN 1.
SQ SEQUENCE 727 AA; 75811 MW; 96DF9B664873985D CRC64;

Query Match 32.9%; Score 54; DB 2; Length 727;
Best Local Similarity 52.4%; Pred. No. 20;
Matches 11; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 9 GPAACGCHGHLILGNPKNSASV 29
Db 568 GGAACGSHLLVTPGKGTVRV 588
|||||
| | | | |

RESULT 9

```

```

Q91LD3
ID Q91LD3 PRELIMINARY; PRT; 312 AA.
AC Q91LD3;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical protein SC01549.
GN SC01549 OR SC111.05C.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seyer K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT *Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).;
RL Nature 417:141-147(2002).
DR EMBL; AL157953; CAB76068.1; -.
KW Hypothetical protein.
SQ SEQUENCE 312 AA; 32324 MW; A9E75F8C1DCE1C9F CRC64;

Query Match 32.3%; Score 53; DB 16; Length 312;
Best Local Similarity 76.9%; Pred. No. 12;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 9 GPAACHGHLHG 21
DQ 179 GPAACHGHLHG 191

RESULT 10
Q9A9Q4
ID Q9A9Q4 PRELIMINARY; PRT; 499 AA.
AC Q9A9Q4;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein CC0920.
GN CC0920.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
OC Caulobacter.
OX NCBI_TaxID=155892;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19089 / CB15;
RX MEDLINE=21173698; PubMed=11259647;
RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Hart D.H.,
RA Klotz J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
RA Utterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RT *Complete genome sequence of Caulobacter crescentus.;
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
DR EMBL; AC005768; AAK22904.1; -.
DR TIGR; CC0920; -.
DR InterPro; IPR001993; Mitoch. carrier.
DR PROSITE; PS00215; MITOCH. CARRIER; UNKNOWN_1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 499 AA; 54933 MW; CCF42A7F833CEAF CRC64;

```

```

Query Match 32.3%; Score 53; DB 16; Length 499;
Best Local Similarity 55.0%; Pred. No. 19;
Matches 11; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 3 DSHLGAGPAACHGHLHG 22
DQ 196 DSHLGAGPAACHGHLHG 215

RESULT 11
Q9LES2
ID Q9LES2 PRELIMINARY; PRT; 478 AA.
AC Q9LES2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical 51.4 kDa protein (AT3G56860/TBM15_190) (UBP1 interacting
DE protein 2a).
GN TBM15_190 OR AT3G56860 OR UBA2A.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Benes V., Wurmbech E., Drzonek H., Ansorge W., Mewes H.W., Rudd S.,
RA Lemcke K., Mayer K.F.X., Quetier F., Salanoubat M.;
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Shinn P.,
RA Tracy S.E., Ban J., Bowser L., Carninci P., Jones T., Kamiya A.,
RA Goldsmith A.D., Hayashizaki Y., Ishida J., Lee J.M., Lin J., Liu S.X.,
RA Karlin-Neumann G., Kawai J., Lam B., Lee C.J., Pham P.K., Quach H.L.,
RA Miranda M., Narusaka M., Nguyen M., Palm C.J., Pham P.K., Quach H.L.,
RA Sakano H., Sakurai T., Satou M., Seki M., Southwick A., Tang C.C.,
RA Toriumi M., Yamada K., Yu G., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RT *Arabidopsis cDNA clones.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Shinn P., Chen H., Cheuk R., Kim C.J., Meyers M.C., Ban J.,
RA Bowser L., Carninci P., Chung M.K., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,
RA Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M., Nguyen M.,
RA Palm C.J., Pham P.K., Quach H.L., Sakano H., Sakurai T., Satou M.,
RA Seki M., Southwick A., Toriumi M., Yamada K., Yu G., Shinozaki K.,
RA Davis R.W., Theologis A., Ecker J.R.;
RT *Arabidopsis cDNA clones.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA Yamada K., Ban J., Chan M.M., Chang C.H., Chang E., Dale J.M.,
RA Tang J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,
RA Deng C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,
RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
RA Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Shinozaki K.,
RA Davis R.W., Ecker J.R., Theologis A.;
RT *Arabidopsis Full Length cDNA Clones.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RA Lambermon M.H.L., Fu Y., Wieczorek Kirk D.A., Dupasquier M.,
RA Filipowicz W., Lorkovic Z.J.;
RT *UBA1 and UBA2, two proteins that interact with UBPI, a

```


RT multifunctional effector of pre-mRNA maturation in plants.*;

RL Mol. Cell. Biol. 0:0-0(2002).
 DR EMBL: AL390921; CAC00749.1; -
 DR EMBL: AY037245; AAK59846.1; -
 DR EMBL: AF367280; AAK56269.1; -
 DR EMBL: AY091040; AAM13861.1; -
 DR EMBL: AJ439404; CAD28672.1; -
 DR InterPro: IPR000504; RNA_rec_mot.
 DR Pfam: PF00076; rtm; 2.
 DR SMART: SM00360; RRM; 2.
 DR PROSITE: PS0102; RRM; 2.
 DR PROSITE: PS00030; RRM_RNP_1; UNKNOWN_1.
 KW Hypothetical protein.
 SQ SEQUENCE 478 AA; 51439 MW; 0CD3653669BE2DA4 CRC64;

Query Match 32.0%; Score 52.5; DB 10; Length 478;
 Best Local Similarity 45.5%; Pred. No. 21;
 Matches 10; Conservative 4; Mismatches 7; Indels 1; Gaps 1;

QY 2 KDHLGAGPAAACGHLLGNP 23
 DB 346 RNDNGYGPFGG-HGLMAGNP 366

RESULT 12

Q94B29 ID Q94B29 PRELIMINARY; PRT; 478 AA.
 AC Q94B29
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Hypothetical 51.4 kDa protein.
 GN TM6_190 OR AT355860.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Nguyen M., Karlin-Neumann G., Southwick A., Lam B., Miranda M.,
 RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
 RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
 RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
 RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
 RA Ecker J., Theologis A., Davis R.W.;
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M.,
 RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
 RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
 RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
 RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
 RA Ecker J., Theologis A., Davis R.W.;
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.

QY 2 KDHLGAGPAAACGHLLGNP 23
 DB 346 RNDNGYGPFGG-HGLMAGNP 366

Query Match 32.0%; Score 52.5; DB 10; Length 478;
 Best Local Similarity 45.5%; Pred. No. 21;
 Matches 10; Conservative 4; Mismatches 7; Indels 1; Gaps 1;

QY 2 KDHLGAGPAAACGHLLGNP 23
 DB 346 RNDNGYGPFGG-HGLMAGNP 366

RESULT 13

Q97FK0 ID Q97FK0 PRELIMINARY; PRT; 408 AA.
 AC Q97FK0
 DT 01-OCT-2001 (TrEMBLrel. 18, Created)
 DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE DNA repair exonuclease.
 GN CAC2737.
 OS Clostridium acetobutylicum.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
 OC Clostridiales; Clostridiaceae; Clostridium.
 OX NCBI_TaxID=1488;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
 RX MEDLINE=21359325; PubMed=11466286;
 RA Noelling J., Bzeon G., Omelchenko M.V., Makarova K.S., zeng Q.,
 RA Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
 RA Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
 RA Bennett G.N., Koonin E.V., Smith D.R.;
 RA "Genome sequence and comparative analysis of the solvent-producing
 bacterium Clostridium acetobutylicum";
 RT J. Bacteriol. 183:4823-4838(2001).
 DR EMBL: AB007771; AAK80683.1; -
 DR InterPro: IPR004843; M-peptidase.
 DR InterPro: IPR004844; S/T-phosphatase.
 DR InterPro: IPR004593; Sbcd.
 DR Pfam: PF00149; Metallophos; 1.
 DR TIGRFAMs: TIGR00619; sbcd; 1.
 KW Exonuclease; Complete proteome.
 SQ SEQUENCE 408 AA; 46682 MW; CAEE3E17CFAE2B39 CRC64;

Query Match 31.7%; Score 52; DB 16; Length 408;
 Best Local Similarity 45.0%; Pred. No. 21;
 Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 6 LGAGPAAACGHLLGNPKN 25
 DB 92 VAAGPLARDHGIIMVGPKS 111

RESULT 14

Q9W619 ID Q9W619 PRELIMINARY; PRT; 141 AA.
 AC Q9W619
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE I22 protein.
 GN I22
 OS Cynops pyrrhogaster (Japanese common newt).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Caudata; Salamandroidae; Salamandridae; Cynops.
 OX NCBI_TaxID=8330;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-REGENERATING LIMB BLASTOMA;
 RT "Expression of the helix-loop-helix type negative regulators of
 RT differentiation during limb regeneration in urodele and anuran";
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB019514; BAA76631.1; -
 DR InterPro: IPR001092; HLH_basic.
 DR Pfam: PF00010; HLH; 1.
 DR SMART: SM00353; HLH; 1.
 DR PROSITE: PS00036; HELIX_LOOP_HELIX; UNKNOWN_1.
 SQ SEQUENCE 141 AA; 15443 MW; C2012A212585C4A5 CRC64;

Query Match 31.1%; Score 51; DB 13; Length 141;
 Best Local Similarity 46.2%; Pred. No. 10;
 Matches 12; Conservative 2; Mismatches 12; Indels 0; Gaps 0;

QY 3 DSHLGAGPAAACHGHLLGNPKNGAS 28
 ||| | | | | | : | |
 Db 83 DSHPGMGAGSGSHIQQRSPSASRS 108

RESULT 15

Q8WZT3 PRELIMINARY; PRT; 187 AA.
 AC Q8WZT3;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Hypothetical 21.3 kDa protein.
 GN B8L21.040.
 OS Neurospora crassa.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariales; Sordariaceae; Neurospora.
 OX NCBI_TaxID=5141;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Schulte U., Align V., Hoheisel J., Brandt P., Fartmann B., Holland R.,
 RA Nyakatura G., Mewes H.W., Mannhaupt G.;
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA German Neurospora genome project;
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL669989; CAD21090.1; -
 KW Hypothetical protein.
 SQ SEQUENCE 187 AA; 21262 MW; 235332C807CC9F24 CRC64;

Query Match 31.1%; Score 51; DB 3; Length 187;
 Best Local Similarity 55.6%; Pred. No. 13;
 Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 IKDSHLGAGPAAACHGHL 18
 | | | | : | | | | | : | | :
 Db 49 IGTSHLSSPAAAHGDM 66

Search completed: June 3, 2003, 19:23:56
 Job time : 9.42683 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2003, 19:18:46 ; Search time 1.76423 Seconds
(without alignments)
728.798 Million cell updates/sec

Title: US-09-938-703-6_COPY_185_215
Perfect score: 164
Sequence: 1 IKDSHGLGAGPAAACHGHLILGNPKNSASYSK 31

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|---------------|--------------------|
| 1 | 56 | 34.1 | 240 | 1 YA02_SCHPO | Q9675 schizosacch |
| 2 | 51.5 | 31.4 | 681 | 1 HMC2_HELPFE | Q9674 helicobacte |
| 3 | 49.5 | 30.2 | 597 | 1 V70K_OYMW | P20130 ononis yell |
| 4 | 48.5 | 29.6 | 222 | 1 NADD_XYLF | Q9675 xyella fas |
| 5 | 48 | 29.3 | 358 | 1 INCL_ECOLI | Q5212 escherichia |
| 6 | 47.5 | 29.0 | 617 | 1 IF2P_AERPE | Q9673 aeropyrum p |
| 7 | 47.5 | 29.0 | 660 | 1 YFEG_ECO57 | Q8423 escherichia |
| 8 | 47.5 | 29.0 | 660 | 1 YFEG_ECOLI | P7798 escherichia |
| 9 | 47.5 | 29.0 | 2319 | 1 YFEG_MOUSE | Q6194 mus musculu |
| 10 | 47.5 | 29.0 | 2359 | 1 CCAH_RAT | Q9660 rattus norv |
| 11 | 47 | 28.7 | 830 | 1 BGIS_BUTFI | P16084 butyrivibri |
| 12 | 46.5 | 28.4 | 516 | 1 CP23_HORVU | P52711 hordeum vul |
| 13 | 46.5 | 28.4 | 1059 | 1 Y379_HUMAN | Q15084 homo sapien |
| 14 | 46 | 28.0 | 211 | 1 CMAA_PARDE | P5218 paracoccus |
| 15 | 46 | 28.0 | 221 | 1 THIE_PASNU | P57930 pasteurella |
| 16 | 46 | 28.0 | 461 | 1 BENA_ACICA | P07769 acinetobact |
| 17 | 46 | 28.0 | 478 | 1 VP26_DROME | Q9452 drosophila |
| 18 | 46 | 28.0 | 650 | 1 Y481_HUMAN | Q75069 homo sapien |
| 19 | 45.5 | 27.7 | 717 | 1 NFCL_MOUSE | Q88942 mus musculu |
| 20 | 45.5 | 27.7 | 816 | 1 QAIF_NEUCR | P11638 neurospora |
| 21 | 45.5 | 27.7 | 1011 | 1 HDA9_HUMAN | Q9400 homo sapien |
| 22 | 45.5 | 27.7 | 1733 | 1 VNVA_PVYKA | P33485 pseudorabie |
| 23 | 45.5 | 27.7 | 2351 | 1 FAS8_HUMAN | P00451 homo sapien |
| 24 | 45 | 27.4 | 376 | 1 HMEV_DROME | P06602 drosophila |
| 25 | 45 | 27.4 | 815 | 1 CC53_YEAST | Q12018 saccharomyc |
| 26 | 45 | 27.4 | 1139 | 1 INAL_CABEL | Q03600 caenorhabdi |
| 27 | 44.5 | 27.1 | 660 | 1 YFPG_SALTY | Q82540 salmonella |
| 28 | 44.5 | 27.1 | 660 | 1 YFPG_SALTY | Q52325 salmonella |
| 29 | 44.5 | 27.1 | 661 | 1 TKT_TREPA | Q83571 treponema p |
| 30 | 44.5 | 27.1 | 670 | 1 TBUD_BURPI | Q01551 burkholderi |
| 31 | 44.5 | 27.1 | 841 | 1 NEK4_HUMAN | P51957 homo sapien |
| 32 | 44 | 26.8 | 188 | 1 RM32_HUMAN | Q9568 homo sapien |
| 33 | 44 | 26.8 | 260 | 1 GTW1_CHICK | P20135 gallus gall |

34 44 26.8 420 1 RA52_MOUSE
35 44 26.8 664 1 VP2_BPPH6
36 44 26.8 690 1 ATKB_SYNT3
37 44 26.8 987 1 PMFG_CHLMU
38 43.5 26.5 309 1 YBCK_HAEIN
39 43.5 26.5 786 1 AAC_ACTUT
40 43.5 26.5 1127 1 MDW1_YEAST
41 43 26.2 208 1 RISA_BUCAL
42 43 26.2 222 1 YA0D_SCHPO
43 43 26.2 253 1 GS2_HUMAN
44 43 26.2 286 1 RRPP_SYNV
45 43 26.2 353 1 XYNA_PENCH

P43352 mus musculu
P11124 bacterioph
P73867 synechocyst
Q9p145 chlamydia m
P44298 haemophilus
P29958 actinoplan
Q01846 saccharomyc
P57212 buchnera ap
Q10092 schizosacch
P41247 homo sapien
P21299 sonchus yel
P29417 penicillium

ALIGNMENTS

RESULT 1
YA02_SCHPO STANDARD; PRT; 240 AA.
AC Q09675;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein C5H10.02c in chromosome I.
GN SPAC5H10.02C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_taxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajadream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feldwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Haidvoigt J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinovitch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA Weijtens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochaz M., Gallardin C., Tallada V.A., Garzon A., Rhode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerrutti L., Lowe R., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakowski G.V., Ussery D., Barrell B.G., Nurse P.;
The genome sequence of Schizosaccharomyces pombe.;
RL Nature 415:871-880(2002).
CC -|- SIMILARITY: BELONGS TO THE YMR322C/YOR391C/YPL280W (YEAST),
CC SPAC1F7.06/SPAC5H10.02C/SPAC1D3.13 (S.POMBE FAMILY).
CC -|- SIMILARITY: SOME, TO E.COLI YEDU.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

| | | | |
|----------|---------|-----------|----------------------------------|
| TRANSFEM | 73 | 92 | 1 (BY SIMILARITY). |
| DOMAIN | 93 | 102 | EXTRACELLULAR (BY SIMILARITY). |
| TRANSFEM | 103 | 124 | 2 (BY SIMILARITY). |
| DOMAIN | 125 | 131 | CYTOPLASMIC (BY SIMILARITY). |
| TRANSFEM | 132 | 151 | 3 (BY SIMILARITY). |
| DOMAIN | 152 | 154 | EXTRACELLULAR (BY SIMILARITY). |
| TRANSFEM | 155 | 174 | 4 (BY SIMILARITY). |
| DOMAIN | 175 | 308 | CYTOPLASMIC (BY SIMILARITY). |
| TRANSFEM | 309 | 327 | 5 (BY SIMILARITY). |
| DOMAIN | 328 | 332 | EXTRACELLULAR (BY SIMILARITY). |
| TRANSFEM | 333 | 350 | 6 (BY SIMILARITY). |
| DOMAIN | 351 | 630 | CYTOPLASMIC (BY SIMILARITY). |
| TRANSFEM | 631 | 652 | 7 (BY SIMILARITY). |
| DOMAIN | 653 | 660 | EXTRACELLULAR (BY SIMILARITY). |
| TRANSFEM | 661 | 676 | 8 (BY SIMILARITY). |
| DOMAIN | 677 | 681 | CYTOPLASMIC (BY SIMILARITY). |
| TRANSFEM | 1 | 63 | HMA. |
| DOMAIN | 388 | 388 | PHOSPHORYLATION (BY SIMILARITY). |
| TRANSFEM | 11 | 11 | POTENTIAL. |
| DOMAIN | 14 | 14 | POTENTIAL. |
| TRANSFEM | 578 | 578 | MAGNESIUM (BY SIMILARITY). |
| DOMAIN | 582 | 582 | MAGNESIUM (BY SIMILARITY). |
| TRANSFEM | 681 AA; | 73747 MW; | 913834388D4FEC3B CRC64; |
| DOMAIN | 681 AA; | 73747 MW; | 913834388D4FEC3B CRC64; |

Query Match 31.4%; Score 51.5; DB 1; Length 681;
Best Local Similarity 52.6%; Pred. No. 6.7;
Matches 10; Conservative 3; Mismatches 5; Indels 1; Gaps 1;

QY 7 GAGPAAACHGHLL--GNPK 24
DB 453 GLGVQASCHSLHIIAGNDK 471

RESULT 3
V7OK_OYMW
ID V7OK_OYMW STANDARD; PRT; 597 AA.
AC P20130;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-FEB-1991 (Rel. 17, Last annotation update)
DE 66 kDa protein.
OS Ononis yellow mosaic virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Tymovirus.
OX NCBI_TaxID=12153;
RN [1]
RP MEDLINE=90021186; PubMed=2800337;
RT Ding S.W., Keese P., Gibbs A.;
RT "Nucleotide sequence of the ononis yellow mosaic tymovirus genome.";
RL Virology 172:555-563(1989).
CC -!- FUNCTION: NOT KNOWN.
CC -!- SIMILARITY: TO 65 TO 70 kDa PROTEIN FROM OTHER TYMOVIRUSES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: J04375; AAA46795.1; -
CC PIR: J00107; J00107.
CC InterPro: IPR004935; Tymo_45_70kDa.
CC Pfam: PF03251; Tymo_45Kd_70kd; 1.
CC SEQUENCE 597 AA; 65985 MW; 8A0BE4C2A7D329F8 CRC64;
CC
CC Query Match 30.2%; Score 49.5; DB 1; Length 597;
CC Best Local Similarity 37.5%; Pred. No. 12;
CC Matches 12; Conservative 4; Mismatches 9; Indels 7; Gaps 1;
CC
CC QY 4 SHLGAPPAACHGHLL-----GNPKNSAS 28

GN INFE OR APE2374.
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;
OC Desulfurococcaceae; Aeropyrum.
NCBI_TaxID=56636;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN-KI;
RX MEDLINE-99310339; PubMed-10382966;
RA Kwarabayasi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankaei H.,
RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamiya J., Masuda S., Funabashi T., Tanaka T., Kudoh Y.,
RA Yamazaki J., Kushida N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sakoi Y., Kikuchi H.;
RT "Complete genome sequence of an aerobic hyper-thermophilic
RT crenarchaeon, Aeropyrum pernix KI.";
RL DNA Res. 6:83-101(1999).
CC -1- FUNCTION: FUNCTION IN GENERAL TRANSLATION INITIATION BY PROMOTING
CC THE BINDING OF THE FORMYLmethionine-tRNA TO RIBOSOMES. SEEKS TO
CC FUNCTION ALONG WITH EIF-2 (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE IF-2 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to licenset@isb-sib.ch).
CC -----
DR EMBL; AP000064; BA81387.1; -.
DR InterPro: IPR004161; EFTU_D2.
DR InterPro: IPR000795; EF_GTPbind.
DR InterPro: IPR00178; IP2.
DR InterPro: IPR005225; Small_GTP.
DR InterPro: IPR004544; IIF_AIF-2.
DR Pfam; PF00009; GTP_EFTU; 1.
DR Pfam; PF03144; GTP_EFTU_D2; 1.
DR PRINTS; PR00315; ELONGATNFC.
DR TIGRFRAMS; TIGR00231; small_GTP; 1.
DR TIGRFRAMS; TIGR00491; aif-2; 1.
DR PROSITE; PS01176; IP2; FALSE_NEG.
KW Initiation factor; Protein biosynthesis; GTP-binding;
KW Complete proteome.
NP_BIND 23 30 GTP (BY SIMILARITY).
FT NP_BIND 87 91 GTP (BY SIMILARITY).
FT NP_BIND 141 144 GTP (BY SIMILARITY).
SQ SEQUENCE 617 AA; 68520 MW; 493DC149C6028EB3 CRC64;

Query Match 29.3%; Score 48; DB 1; Length 617;
Best Local Similarity 27.6%; Pred. No. 20;
Matches 8; Conservative 8; Mismatches 13; Indels 0; Gaps 0;

QY 1 IKDSHLAGPAAACHGHLHGKPNKNSASV 29
DB 531 LEEAELGRAVAVSTQRELLIGHANEEDI 559

RESULT 7
YFPG_EC057
ID YFPG_EC057 STANDARD; PRT; 660 AA.
AC Q8XDZ3;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Hypothetical protein yfpg.
DN YFPG OR Z3513; OR ECS3143.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
NCBI_TaxID=83334;
[1]

RN SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RX MEDLINE-97426617; PubMed-9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."

SEQUENCE FROM N.A.
MEDLINE-O157:H7 / EDL933 / ATCC 700927;
STRAIN-21074935; PubMed-11206551;
Perma N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
Postfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
Weich R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";
RL Nature 409:529-533(2001).
[2]
RN SEQUENCE FROM N.A.
RC STRAIN-O157:H7 / RMD 0509952;
EX MEDLINE-21156231; PubMed-11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsuka E., Nakayama K., Murata T., Tanaka M., Toke T.,
RA Liida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shibata T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
O157:H7 and genomic comparison with a laboratory strain K-12";
RL DNA Res. 8:11-22(2001).
CC -1- SIMILARITY: BELONGS TO THE EHEC FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to licenset@isb-sib.ch).

EMBL; AE005458; AAC57386.1; -.
EMBL; AP002561; BAB36566.1; -.
InterPro: IPR002376; formyl_transf.
InterPro: IPR004187; Formyl_trans_C.
Pfam; PF00551; formyl_transf; 1.
Pfam; PF02911; formyl_trans_C; 1.
KW Hypothetical protein; Transferase; Methyltransferase;
KW Complete proteome.
SQ SEQUENCE 660 AA; 74209 MW; 9D45B3AF60CF6131 CRC64;

Query Match 29.0%; Score 47.5; DB 1; Length 660;
Best Local Similarity 42.1%; Pred. No. 25;
Matches 8; Conservative 6; Mismatches 4; Indels 1; Gaps 1;

QY 14 CHGHLL-LGNPKNSASVSXK 31
DB 558 CDGEIINIGNPENASIEE 576

RESULT 8
YFPG_EC01
ID YFPG_EC01 STANDARD; PRT; 660 AA.
AC P77398;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein yfpg.
DN YFPG OR B2255.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
NCBI_TaxID=562;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RX MEDLINE-97426617; PubMed-9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."

FT CARBOHYD 1255 1255 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1268 1268 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1273 1273 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1274 1274 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1302 1302 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1316 1316 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1340 1340 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1378 1378 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1797 1797 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 2105 2105 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 2319 AA; 266148 MW; FD054DE051DB2A01 CRC64;

Query Match 29.0%; Score 47.5; DB 1; Length 2319;
 Best Local Similarity 40.0%; Pred. No. 86;
 Matches 12; Conservative 3; Mismatches 12; Indels 3; Gaps 2;

OY 2 KDSHLG-AGPAAACHGILLGNPKNSASVS 30
 DB 1832 RDMHSLGLPILCHANTL--NPAHGQVVS 1859

RESULT 10
 CCAH_RAT
 ID CCAH_RAT STANDARD; PRO; 2359 AA.
 AC Q9P060;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DE Voltage-dependent T-type calcium channel alpha-1H subunit (Cav3.1c).
 GN CACNA1H.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=21264893; PubMed=11073957;
 RA McRory J.E., Santi C.M., Hamming K.S.C., Mezeyova J., Sutton K.G.,
 RA Baillie D.L., Stea A., Snutch T.P.;
 RT "Molecular and functional characterization of a family of rat brain
 T-type calcium channels";
 RL J. Biol. Chem. 276:3999-4011(2001).
 CC - FUNCTION: VOLTAGE-SENSITIVE CALCIUM CHANNELS (VSCC) MEDIATE THE
 ENTRY OF CALCIUM IONS INTO EXCITABLE CELLS AND ARE ALSO INVOLVED
 IN A VARIETY OF CALCIUM-DEPENDENT PROCESSES, INCLUDING MUSCLE
 CONTRACTION, HORMONE OR NEUROTRANSMITTER RELEASE, GENE EXPRESSION,
 CELL MOTILITY, CELL DIVISION AND CELL DEATH. THE ISOFORM ALPHA-1H
 GIVES RISE TO T-TYPE CALCIUM CURRENTS. T-TYPE CALCIUM CHANNELS
 BELONG TO THE "LOW-VOLTAGE ACTIVATED (LVA)" GROUP AND ARE STRONGLY
 BLOCKED BY NICKEL AND MIBEPRADIL. A PARTICULARITY OF THIS TYPE OF
 CHANNELS IS AN OPENING AT QUITE NEGATIVE POTENTIALS. AND A
 VOLTAGE-DEPENDENT INACTIVATION. T-TYPE CHANNELS SERVE PACEMAKING
 FUNCTIONS IN BOTH CENTRAL NEURONS AND CARDIAC NODAL CELLS AND
 SUPPORT CALCIUM SIGNALING IN SECRETORY CELLS AND VASCULAR SMOOTH
 MUSCLE. THEY MAY ALSO BE INVOLVED IN THE MODULATION OF FIRING
 PATTERNS OF NEURONS WHICH IS IMPORTANT FOR INFORMATION PROCESSING
 AS WELL AS IN CELL GROWTH PROCESSES.
 CC - SUBCELLULAR LOCATION: Integral membrane protein.
 CC - TISSUE SPECIFICITY: Expressed in brain.
 CC - DOMAIN: EACH OF THE FOUR INTERNAL REPEATS CONTAINS FIVE
 HYDROPHOBIC TRANSMEMBRANE SEGMENTS (S1, S2, S3, S5, S6) AND ONE
 POSITIVELY CHARGED TRANSMEMBRANE SEGMENT (S4). S4 SEGMENTS
 PROBABLY REPRESENT THE VOLTAGE-SENSOR AND ARE CHARACTERIZED BY A
 SERIES OF POSITIVELY CHARGED AMINO ACIDS AT EVERY THIRD POSITION.
 CC - PFM: IN RESPONSE TO RAISING OF INTRACELLULAR CALCIUM, THE T-TYPE
 CHANNELS ARE ACTIVATED BY CAM-KINASE II.
 CC - SIMILARITY: BELONGS TO THE CALCIUM CHANNEL ALPHA-1 SUBUNITS
 FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its

use by non-profit Institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 EMBL: AF290213; AAC35187.1; -
 InterPro: IPR001682; Ca/Na_pore.
 InterPro: IPR002111; Cat_channel_TrpL.
 InterPro: IPR002077; Ca_channel.
 InterPro: IPR000636; M-channel_nlg.
 InterPro: IPR003915; PKD.2.
 Pfam: PF00520; ion.trans. 4.
 PRINTS: PR00167; CACHANNEL.
 PRINTS: PR01433; POLYCYSTIN2.
 PRINTS: PR01629; TVDCCALPHA1.
 Ionic channel; Transmembrane; Ion transport; Voltage-gated channel;
 Calcium channel; Glycoprotein; Repeat; Multigene family;
 Phosphorylation.
 FT REPEAT 87 422
 FT REPEAT 776 1015
 FT REPEAT 1292 1569
 FT REPEAT 1607 1868
 FT DOMAIN 1 100
 FT TRANSMEM 101 119
 FT DOMAIN 120 139
 FT TRANSMEM 140 160
 FT DOMAIN 161 169
 FT TRANSMEM 170 184
 FT DOMAIN 185 193
 FT TRANSMEM 194 212
 FT DOMAIN 213 232
 FT TRANSMEM 233 253
 FT DOMAIN 254 394
 FT TRANSMEM 395 419
 FT DOMAIN 420 790
 FT TRANSMEM 791 811
 FT DOMAIN 812 824
 FT TRANSMEM 825 846
 FT DOMAIN 847 852
 FT TRANSMEM 853 871
 FT DOMAIN 872 879
 FT TRANSMEM 880 903
 FT DOMAIN 904 914
 FT TRANSMEM 915 935
 FT DOMAIN 936 987
 FT TRANSMEM 988 1022
 FT DOMAIN 1013 1301
 FT TRANSMEM 1302 1324
 FT DOMAIN 1325 1342
 FT TRANSMEM 1343 1363
 FT DOMAIN 1364 1373
 FT TRANSMEM 1374 1393
 FT DOMAIN 1394 1407
 FT TRANSMEM 1408 1429
 FT DOMAIN 1430 1439
 FT TRANSMEM 1440 1463
 FT DOMAIN 1464 1540
 FT TRANSMEM 1541 1566
 FT DOMAIN 1567 1621
 FT TRANSMEM 1622 1642
 FT DOMAIN 1643 1656
 FT TRANSMEM 1657 1678
 FT DOMAIN 1679 1685
 FT TRANSMEM 1686 1704
 FT DOMAIN 1705 1718
 FT TRANSMEM 1719 1742
 FT DOMAIN 1743 1756
 FT TRANSMEM 1757 1777
 FT DOMAIN 1778 1840
 FT TRANSMEM 1841 1868
 FT DOMAIN 1869 2359
 FT POLY-HIS.
 FT POLY-ARG.
 FT 1594

| | | | | |
|-----------------------|--------|---------------|-------|-------------|
| Query Match | 28.7%; | Score 47; | DB 1; | Length 830; |
| Best Local Similarity | 37.0%; | Pred. No. 38; | | |

N-LINKED (GLCNAC. . .) (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).

```
FT CARBOHYD 380 380 N-LINKED (GLCNAC... ) (POTENTIAL).
SQ SEQUENCE 516 AA; 55913 MW; D4IAAIC56CF8D355 CRC64;

Query Match
Best Local Similarity 28.4%; Score 46.5; DB 1; Length 516;
Matches 13; Conservative 3; Mismatches 7; Indels 9; Gaps 1;

QY 6 LGAGPAAACHG-----HLLGNPKNSAS 28
      |||||
DB 19 LGAGPAAAKARRTRQGDYINRLRGSPSSRAS 50

RESULT 13
Y379_HUMAN
ID Y379_HUMAN STANDARD; PRT; 1059 AA.
AC O15084;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Hypothetical protein KIAA0379 (Fragment).
GN KIAA0379.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=97349984; PubMed=9205841;
RA Nagase T., Ishikawa K.-I., Nakajima D., Ohira M., Seki N.,
RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. VII.
RT The complete sequences of 100 new cDNA clones from brain which can
RT code for large proteins in vitro.";
RL DNA Res. 4:141-150(1997).
CC -!- SIMILARITY: CONTAINS AT LEAST 27 ANK REPEATS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB002377; BAA20833.2; -
CC DR HSP; P80144; 2MYO.
CC DR InterPro: IPR002110; ANK.
CC DR Pfam; PF00023; ank; 28.
CC DR SMART; SM00248; ANK; 26.
CC DR PROSITE; PS50088; ANK_REPEAT; 24.
CC DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
CC KW Hypothetical protein; Repeat; ANK repeat.
FT NON_TER 1
FT REPEAT 46 75 ANK 1.
FT REPEAT 79 108 ANK 2.
FT REPEAT 112 141 ANK 3.
FT REPEAT 145 174 ANK 4.
FT REPEAT 178 207 ANK 5.
FT REPEAT 211 240 ANK 6.
FT REPEAT 244 273 ANK 7.
FT REPEAT 277 307 ANK 8.
FT REPEAT 311 340 ANK 9.
FT REPEAT 344 373 ANK 10.
FT REPEAT 377 406 ANK 11.
FT REPEAT 410 439 ANK 12.
FT REPEAT 443 472 ANK 13.
FT REPEAT 476 506 ANK 14.
FT REPEAT 510 540 ANK 15.
FT REPEAT 555 584 ANK 16.
FT REPEAT 588 617 ANK 17.
FT REPEAT 622 651 ANK 18.
FT REPEAT 658 687 ANK 19.
```

```
FT REPEAT 691 720 ANK 20.
FT REPEAT 724 753 ANK 21.
FT REPEAT 761 790 ANK 22.
FT REPEAT 793 823 ANK 23.
FT REPEAT 828 857 ANK 24.
FT REPEAT 861 891 ANK 25.
FT REPEAT 895 924 ANK 26.
FT REPEAT 931 960 ANK 27.
SQ SEQUENCE 1059 AA; 113465 MW; C1F55E6CFB494770 CRC64;

Query Match
Best Local Similarity 28.4%; Score 46.5; DB 1; Length 1059;
Matches 15; Conservative 6; Mismatches 8; Indels 9; Gaps 3;

QY 1 IKDHLGAGP--AAACHG-----LLGN--PKNSASV 29
      :||| | | | | | | | | | | | | | | | |
DB 617 VKDYILKRTPIHAAATNGHSECLRLIGNAEQNQAVDI 654

RESULT 14
CCMA_PARDE
ID CCMA_PARDE STANDARD; PRT; 211 AA.
AC P52218;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Heme exporter protein A (Cytochrome c-type biogenesis ATP-binding
DE protein ccma).
DE CCMA.
OS Paracoccus denitrificans.
OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;
OC Paracoccus.
OX NCBI_TaxID=266;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Pd 1222;
RX MEDLINE=97195802; PubMed=9043133;
RA Page D., Pearce D.A., Norris H.A., Ferguson S.J.;
RT "The Paracoccus denitrificans ccma, B and C genes: cloning and
RT sequencing, and analysis of the potential of their products to form a
RT haem or apo-c-type cytochrome transporter.";
RL Microbiology 143:563-576(1997).
CC -!- FUNCTION: REQUIRED FOR THE EXPORT OF HEME TO THE PERIPLASM FOR THE
CC BIODEGRADATION OF C-TYPE CYTOCHROMES (BY SIMILARITY). (Potential).
CC -!- SUBCELLULAR LOCATION: Inner membrane-associated (Potential).
CC -!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z71971; CAA96502.1; -
CC DR InterPro: IPR003593; AAA_ATPase.
CC DR InterPro: IPR003439; ABC_transportr.
CC DR Pfam; PF00005; ABC_tran; 1.
CC DR PRODOM; PD000006; ABC_transportr; 1.
CC DR SMART; SM00382; AAA; 1.
CC DR TIGRFAMs; TIGR01189; CCMA; 1.
CC DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
CC KW Cytochrome c-type biogenesis; Transport; ATP-binding; Inner membrane.
FT NP_BIND 36 43 ATP (POTENTIAL).
SQ SEQUENCE 211 AA; 21978 MW; 3C40E54C1B37CEFC CRC64;

Query Match
Best Local Similarity 28.0%; Score 46; DB 1; Length 211;
Matches 10; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 4 SHLGAGPAAACHGHLGNPK 24
      :||| | | | | | | | | | | | | | | | |
```

```

Db      163 AHIGGGAIVATHIDGLPE 183

RESULT 15
THIE_PASMO
ID THIE_PASMO STANDARD; PRT; 221 AA.
AC P57930;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Thiamine-phosphate pyrophosphorylase (EC 2.5.1.3) (TMP
DE pyrophosphorylase) (TMP-PPase) (Thiamine-phosphate synthase).
GN THIE OR PM1260.
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Pasteurella.
OX NCBI_TaxID=747;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Pw70;
RX MEDLINE=21145866; PubMed=11248100;
RA May B.J., Zhang Q., Li L.L., Faustian M.L., Whittam T.S., Kapur V.;
RF "Complete genomic sequence of Pasteurella multocida Pw70.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
CC -1- FUNCTION: CONDENSES 4-METHYL-5-(BETA-HYDROXYETHYL)-THIAZOLE
CC MONOPHOSPHATE (THZ-P) AND 4-AMINO-5-HYDROXYMETHYL PYRIMIDINE
CC PYROPHOSPHATE (HMP-PP) TO FORM THIAMINE MONOPHOSPHATE (TMP) (BY
CC SIMILARITY).
CC -1- CATALYTIC ACTIVITY: 2-methyl-4-amino-5-hydroxymethylpyrimidine
CC diphosphate + 4-4-methyl-5-(2-phosphonooxyethyl)-thiazole =
CC diphosphate + thiamine monophosphate.
CC -1- PATHWAY: Thiamine biosynthesis.
CC -1- SIMILARITY: BELONGS TO THE TMP-PPASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE006165; AAK03344.1; -.
DR HSSP; P39594; 2TPS.
DR InterPro; IPR003733; TMP_synthase.
DR Pfam; PF02581; TMP-TENI; 1.
DR TIGRams; TIGR00693; thie; 1.
KW Thiamine biosynthesis; transferase; Complete proteome.
SQ SEQUENCE 221 AA; 23736 MW; 2207E7CF5E62134 CRC64;

Query Match 28.0%; Score 46; DB 1; Length 221;
Best Local Similarity 35.7%; Pred. No. 14;
Matches 10; Conservative 3; Mismatches 9; Indels 6; Gaps 1;

QY 5 HLG-----AGPAACHGHLIGPKNS 26
|:| | | | | | | | | |
Db 97 RVGQTDMAVDVAALCHSCFTGTSVNT 124

Search completed: June 3, 2003, 19:22:57
Job time : 2.76423 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2003, 19:24:01 ; Search time 7.43496 Seconds
(without alignments)
422.051 Million cell updates/sec

Title: US-09-938-703-6_COPY_185_215

Perfect score: 164

Sequence: 1 IKDHLGAGPAAACHGILLGNPKNSASVSK 31

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 383519 seqs, 101223694 residues

Total number of hits satisfying chosen parameters: 383519

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA:

1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|---------------------|
| 1 | 164 | 100.0 | 34 | 10 | US-09-938-719-13 |
| 2 | 164 | 100.0 | 34 | 10 | US-09-939-226-13 |
| 3 | 164 | 100.0 | 34 | 10 | US-09-938-703-13 |
| 4 | 164 | 100.0 | 215 | 10 | US-09-938-719-6 |
| 5 | 164 | 100.0 | 215 | 10 | US-09-939-226-6 |
| 6 | 164 | 100.0 | 215 | 10 | US-09-938-703-6 |
| 7 | 55 | 33.5 | 84 | 10 | US-09-864-761-43094 |
| 8 | 48 | 29.3 | 297 | 10 | US-09-938-330-4 |
| 9 | 48 | 29.3 | 451 | 10 | US-09-938-330-2 |
| 10 | 48 | 29.3 | 486 | 10 | US-09-938-330-6 |
| 11 | 48 | 29.3 | 1213 | 10 | US-09-938-330-14 |
| 12 | 48 | 29.3 | 1216 | 10 | US-09-938-330-12 |
| 13 | 48 | 29.3 | 1219 | 10 | US-09-938-330-10 |
| 14 | 48 | 29.3 | 1222 | 10 | US-09-938-330-8 |
| 15 | 48 | 29.3 | 1223 | 9 | US-10-205-368-4 |
| 16 | 48 | 29.3 | 1223 | 12 | US-10-014-070-2 |
| 17 | 48 | 29.3 | 1232 | 10 | US-09-938-330-18 |
| 18 | 48 | 29.3 | 1233 | 9 | US-10-205-368-2 |
| 19 | 48 | 29.3 | 1235 | 10 | US-09-938-330-16 |

| | | | | | |
|----|------|------|------|----|--------------------|
| 20 | 48 | 29.3 | 1249 | 10 | US-09-938-330-22 |
| 21 | 48 | 29.3 | 1252 | 10 | US-09-938-330-20 |
| 22 | 48 | 29.3 | 1360 | 9 | US-10-205-368-5 |
| 23 | 47.5 | 29.0 | 124 | 10 | US-09-814-122-62 |
| 24 | 47.5 | 29.0 | 2319 | 9 | US-10-187-319-6 |
| 25 | 47 | 28.7 | 170 | 10 | US-09-925-300-1726 |
| 26 | 47 | 28.7 | 391 | 12 | US-10-062-254-212 |
| 27 | 46 | 28.0 | 190 | 9 | US-09-738-626-3578 |
| 28 | 46 | 28.0 | 352 | 10 | US-09-860-351-2 |
| 29 | 46 | 28.0 | 372 | 9 | US-09-973-941-4 |
| 30 | 46 | 28.0 | 372 | 9 | US-09-986-992-2 |
| 31 | 46 | 28.0 | 372 | 10 | US-09-973-963-4 |
| 32 | 46 | 28.0 | 372 | 10 | US-09-973-064-4 |
| 33 | 46 | 28.0 | 372 | 10 | US-09-973-077-4 |
| 34 | 46 | 28.0 | 372 | 10 | US-09-973-063-4 |
| 35 | 46 | 28.0 | 372 | 10 | US-09-973-964-4 |
| 36 | 46 | 28.0 | 372 | 10 | US-09-975-072-4 |
| 37 | 46 | 28.0 | 372 | 10 | US-09-972-038-4 |
| 38 | 46 | 28.0 | 372 | 10 | US-09-972-757-4 |
| 39 | 46 | 28.0 | 372 | 10 | US-09-973-965-4 |
| 40 | 46 | 28.0 | 373 | 10 | US-09-925-300-1655 |
| 41 | 45.5 | 27.7 | 495 | 9 | US-10-072-094-95 |
| 42 | 45.5 | 27.7 | 546 | 9 | US-10-173-539-10 |
| 43 | 45.5 | 27.7 | 590 | 9 | US-10-072-094-99 |
| 44 | 45.5 | 27.7 | 590 | 9 | US-10-173-539-11 |
| 45 | 45.5 | 27.7 | 780 | 9 | US-10-072-094-93 |

ALIGNMENTS

RESULT 1

US-09-938-719-13
; Sequence 13, Application US/09938719
; Patent No. US20020106742A1
; GENERAL INFORMATION:
; APPLICANT: SAMSON, MICHEL
; PARMENTIER, MARC
; VASSART, GILBERT
; LIBERT, FREDERICK

TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR
AND NUCLEIC ACID MOLECULES ENCODING SAID RECEPTOR

NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESSES:

ADDRESSEE: Knobbe, Martens, Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: CA
COUNTRY: U.S.A.
ZIP: 92660

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/938,719
FILING DATE: 24-Aug-2001
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/626,939
FILING DATE: 27-JULY-2000
ATTORNEY/AGENT INFORMATION:
NAME: Altman, Daniel E
REGISTRATION NUMBER: 34,115
REFERENCE/DOCKET NUMBER: <Unknown>

INFORMATION FOR SEQ ID NO: 13:

LENGTH: 34 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 13:

Sequence 22, Appl
Sequence 20, Appl
Sequence 5, Appl
Sequence 62, Appl
GENERAL INFORMA
Sequence 1726, Ap
Sequence 212, App
Sequence 3578, Ap
Sequence 2, Appl
Sequence 4, Appl
Sequence 4, Appl
Sequence 4, Appl
Sequence 4, Appl
Sequence 4, Appl
Sequence 4, Appl
Sequence 4, Appl
Sequence 1655, Ap
Sequence 95, Appl
Sequence 10, Appl
Sequence 99, Appl
Sequence 11, Appl
Sequence 93, Appl

US-09-938-719-13

Query Match 100.0%; Score 164; DB 10; Length 34;
Best Local Similarity 100.0%; Pred. No. 5.9e-17;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKDHLGAGPAAACHGHLILGNPKNSASVSK 31
DB 4 IKDHLGAGPAAACHGHLILGNPKNSASVSK 34

RESULT 2

US-09-939-226-13

; Sequence 13, Application US/09939226

; Patent No. US20020110805A1

; GENERAL INFORMATION:

; APPLICANT: SAMSON, MICHEL

; PARENTIER, MARC

; VASSART, GILBERT

; LIBERT, FREDERICK

; TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR
AND NUCLEIC ACID MOLECULES ENCODING SAID RECEPTOR

; NUMBER OF SEQUENCES: 17

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Knobbe, Martens, Olson & Bear

; STREET: 620 Newport Center Drive 16th Floor

; CITY: Newport Beach

; STATE: CA

; COUNTRY: U.S.A.

; ZIP: 92660

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/939,226

; FILING DATE: 24-Aug-2001

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 09/626,939

; FILING DATE: 2000-07-27

; ATTORNEY/AGENT INFORMATION:

; NAME: Altman, Daniel E

; REGISTRATION NUMBER: 34,115

; REFERENCE/DOCKET NUMBER: <Unknown>

; INFORMATION FOR SEQ ID NO: 13:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 34 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; SEQUENCE DESCRIPTION: SEQ ID NO: 13:

US-09-939-226-13

Query Match 100.0%; Score 164; DB 10; Length 34;
Best Local Similarity 100.0%; Pred. No. 5.9e-17;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKDHLGAGPAAACHGHLILGNPKNSASVSK 31
DB 4 IKDHLGAGPAAACHGHLILGNPKNSASVSK 34

RESULT 3

US-09-938-703-13

; Sequence 13, Application US/09938703

; Patent No. US20020110870A1

; GENERAL INFORMATION:

; APPLICANT: SAMSON, MICHEL

; PARENTIER, MARC

; VASSART, GILBERT

; LIBERT, FREDERICK

; TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR
AND NUCLEIC ACID MOLECULES ENCODING SAID RECEPTOR

; NUMBER OF SEQUENCES: 17

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Knobbe, Martens, Olson & Bear

; STREET: 620 Newport Center Drive 16th Floor

; CITY: Newport Beach

; STATE: CA

; COUNTRY: U.S.A.

; ZIP: 92660

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/938,703

; FILING DATE: 24-Aug-2001

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 09/626,939

; FILING DATE: 2000-07-27

; ATTORNEY/AGENT INFORMATION:

; NAME: Altman, Daniel E

; REGISTRATION NUMBER: 34,115

; REFERENCE/DOCKET NUMBER: <Unknown>

; INFORMATION FOR SEQ ID NO: 13:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 34 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; SEQUENCE DESCRIPTION: SEQ ID NO: 13:

US-09-938-703-13

Query Match 100.0%; Score 164; DB 10; Length 34;
Best Local Similarity 100.0%; Pred. No. 5.9e-17;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKDHLGAGPAAACHGHLILGNPKNSASVSK 31
DB 4 IKDHLGAGPAAACHGHLILGNPKNSASVSK 34

RESULT 4

US-09-938-719-6

; Sequence 6, Application US/09938719

; Patent No. US20020106742A1

; GENERAL INFORMATION:

; APPLICANT: SAMSON, MICHEL

; PARENTIER, MARC

; VASSART, GILBERT

; LIBERT, FREDERICK

; TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR

AND NUCLEIC ACID MOLECULES ENCODING SAID RECEPTOR

; NUMBER OF SEQUENCES: 17

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Knobbe, Martens, Olson & Bear

; STREET: 620 Newport Center Drive 16th Floor

; CITY: Newport Beach

; STATE: CA

; COUNTRY: U.S.A.

; ZIP: 92660

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/938,719

; FILING DATE: 24-Aug-2001

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

```

; APPLICATION NUMBER: 09/626,939
; FILING DATE: 27-JULY-2000
; ATTORNEY/AGENT INFORMATION:
; NAME: Altman, Daniel E
; REGISTRATION NUMBER: 34,115
; REFERENCE/DOCKET NUMBER: <Unknown>
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 215 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-938-719-6

Query Match 100.0%; Score 164; DB 10; Length 215;
Best Local Similarity 100.0%; Pred. No. 5e-16;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKDHLGAGPAAACHGHLGLGNPKNSASVSK 31
Db 185 IKDHLGAGPAAACHGHLGLGNPKNSASVSK 215

RESULT 5
US-09-939-226-6
; Sequence 6, Application US/09939226
; Patent No. US20020110805A1
; GENERAL INFORMATION:
; APPLICANT: SAMSON, MICHEL
; PARMENTIER, MARC
; VASSART, GILBERT
; LIBERT, FREDERICK
; TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/939,226
; FILING DATE: 24-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/626,939
; FILING DATE: 2000-07-27
; ATTORNEY/AGENT INFORMATION:
; NAME: Altman, Daniel E
; REGISTRATION NUMBER: 34,115
; REFERENCE/DOCKET NUMBER: <Unknown>
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 215 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-939-226-6

Query Match 100.0%; Score 164; DB 10; Length 215;
Best Local Similarity 100.0%; Pred. No. 5e-16;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKDHLGAGPAAACHGHLGLGNPKNSASVSK 31
Db 185 IKDHLGAGPAAACHGHLGLGNPKNSASVSK 215

RESULT 6
US-09-938-703-6
; Sequence 6, Application US/09938703
; Patent No. US20020110870A1
; GENERAL INFORMATION:
; APPLICANT: SAMSON, MICHEL
; PARMENTIER, MARC
; VASSART, GILBERT
; LIBERT, FREDERICK
; TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/938,703
; FILING DATE: 24-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/626,939
; FILING DATE: 2000-07-27
; ATTORNEY/AGENT INFORMATION:
; NAME: Altman, Daniel E
; REGISTRATION NUMBER: 34,115
; REFERENCE/DOCKET NUMBER: <Unknown>
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 215 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-938-703-6

Query Match 100.0%; Score 164; DB 10; Length 215;
Best Local Similarity 100.0%; Pred. No. 5e-16;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKDHLGAGPAAACHGHLGLGNPKNSASVSK 31
Db 185 IKDHLGAGPAAACHGHLGLGNPKNSASVSK 215

RESULT 7
US-09-864-761-43094
; Sequence 43094, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aemica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
```

```

; APPLICATION NUMBER: 09/626,939
; FILING DATE: 27-JULY-2000
; ATTORNEY/AGENT INFORMATION:
; NAME: Altman, Daniel E
; REGISTRATION NUMBER: 34,115
; REFERENCE/DOCKET NUMBER: <Unknown>
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 215 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-938-719-6

Query Match 100.0%; Score 164; DB 10; Length 215;
Best Local Similarity 100.0%; Pred. No. 5e-16;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKDHLGAGPAAACHGHLGLGNPKNSASVSK 31
Db 185 IKDHLGAGPAAACHGHLGLGNPKNSASVSK 215

RESULT 5
US-09-939-226-6
; Sequence 6, Application US/09939226
; Patent No. US20020110805A1
; GENERAL INFORMATION:
; APPLICANT: SAMSON, MICHEL
; PARMENTIER, MARC
; VASSART, GILBERT
; LIBERT, FREDERICK
; TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/939,226
; FILING DATE: 24-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/626,939
; FILING DATE: 2000-07-27
; ATTORNEY/AGENT INFORMATION:
; NAME: Altman, Daniel E
; REGISTRATION NUMBER: 34,115
; REFERENCE/DOCKET NUMBER: <Unknown>
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 215 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-939-226-6

Query Match 100.0%; Score 164; DB 10; Length 215;
Best Local Similarity 100.0%; Pred. No. 5e-16;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKDHLGAGPAAACHGHLGLGNPKNSASVSK 31
Db 185 IKDHLGAGPAAACHGHLGLGNPKNSASVSK 215

RESULT 6
US-09-938-703-6
; Sequence 6, Application US/09938703
; Patent No. US20020110870A1
; GENERAL INFORMATION:
; APPLICANT: SAMSON, MICHEL
; PARMENTIER, MARC
; VASSART, GILBERT
; LIBERT, FREDERICK
; TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/938,703
; FILING DATE: 24-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/626,939
; FILING DATE: 2000-07-27
; ATTORNEY/AGENT INFORMATION:
; NAME: Altman, Daniel E
; REGISTRATION NUMBER: 34,115
; REFERENCE/DOCKET NUMBER: <Unknown>
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 215 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-938-703-6

Query Match 100.0%; Score 164; DB 10; Length 215;
Best Local Similarity 100.0%; Pred. No. 5e-16;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKDHLGAGPAAACHGHLGLGNPKNSASVSK 31
Db 185 IKDHLGAGPAAACHGHLGLGNPKNSASVSK 215

RESULT 7
US-09-864-761-43094
; Sequence 43094, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aemica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
```

```

; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 43094
; LENGTH: 84
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL096816.12
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN LONG, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.93
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.92
; OTHER INFORMATION: EST_HUMAN HIT: A1896698.1, EVALU6 7.00e-28
; US-09-864-761-43094

Query Match      33.5%; Score 55; DB 10; Length 84;
Best Local Similarity 50.0%; Pred. No. 0.98;
Matches 12; Conservative 1; Mismatches 5; Indels 6; Gaps 1;

QY      3  DSHLGAGPAAACHGHLILGNPKNS 26
        ||||| |||||
Db      48  DSHLGGGPAATA-----GGPRTS 65

RESULT 8
US-09-938-330-4
; Sequence 4, Application US/09938330
; Patent No. US20020115838A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Hilbun, Erin
; APPLICANT: Scoville, John
; APPLICANT: Fiddle, Carl Johan
; APPLICANT: Hu, Yi
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. US20020115838A1 Human Proteases and Polynucleotides Encod
; FILE REFERENCE: LEX-0237-USA
; CURRENT APPLICATION NUMBER: US/09/938,330
; CURRENT FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 2001-08-22
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: US 60/227,104
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: US 60/233,796
; PRIOR FILING DATE: 2000-09-19
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 297
; TYPE: PRT
; ORGANISM: homo sapiens
; US-09-938-330-4

Query Match      29.3%; Score 48; DB 10; Length 297;
Best Local Similarity 45.0%; Pred. No. 43;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY      4  SHLGAGPAAACHGHLILGNP 23
        ||: ||||| |:: |
Db      54  SHVSGPAAASAGSMVVDTP 73

RESULT 9
US-09-938-330-2
; Sequence 2, Application US/09938330
; Patent No. US20020115838A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Hilbun, Erin
; APPLICANT: Scoville, John
; APPLICANT: Fiddle, Carl Johan
; APPLICANT: Hu, Yi
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. US20020115838A1 Human Proteases and Polynucleotides Enco
; FILE REFERENCE: LEX-0237-USA
; CURRENT APPLICATION NUMBER: US/09/938,330
; CURRENT FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: US 60/227,104
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: US 60/233,796
; PRIOR FILING DATE: 2000-09-19
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 451
; TYPE: PRT
; ORGANISM: homo sapiens
; US-09-938-330-2

Query Match      29.3%; Score 48; DB 10; Length 451;
Best Local Similarity 45.0%; Pred. No. 70;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY      4  SHLGAGPAAACHGHLILGNP 23
        ||: ||||| |:: |
Db      54  SHVSGPAAASAGSMVVDTP 73

RESULT 10
US-09-938-330-6
; Sequence 6, Application US/09938330
; Patent No. US20020115838A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Hilbun, Erin
; APPLICANT: Scoville, John
; APPLICANT: Fiddle, Carl Johan
; APPLICANT: Hu, Yi
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. US20020115838A1 Human Proteases and Polynucleotides Enco
; FILE REFERENCE: LEX-0237-USA
; CURRENT APPLICATION NUMBER: US/09/938,330
; CURRENT FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 2001-08-22
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: US 60/227,104
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: US 60/233,796
; PRIOR FILING DATE: 2000-09-19
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 297
; TYPE: PRT
; ORGANISM: homo sapiens
; US-09-938-330-4
```

```

; PRIOR APPLICATION NUMBER: US 60/227,104
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: US 60/233,796
; PRIOR FILING DATE: 2000-09-19
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 297
; TYPE: PRT
; ORGANISM: homo sapiens
; US-09-938-330-4

Query Match      29.3%; Score 48; DB 10; Length 297;
Best Local Similarity 45.0%; Pred. No. 43;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY      4  SHLGAGPAAACHGHLILGNP 23
        ||: ||||| |:: |
Db      54  SHVSGPAAASAGSMVVDTP 73

RESULT 9
US-09-938-330-2
; Sequence 2, Application US/09938330
; Patent No. US20020115838A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Hilbun, Erin
; APPLICANT: Scoville, John
; APPLICANT: Fiddle, Carl Johan
; APPLICANT: Hu, Yi
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. US20020115838A1 Human Proteases and Polynucleotides Enco
; FILE REFERENCE: LEX-0237-USA
; CURRENT APPLICATION NUMBER: US/09/938,330
; CURRENT FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: US 60/227,104
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: US 60/233,796
; PRIOR FILING DATE: 2000-09-19
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 451
; TYPE: PRT
; ORGANISM: homo sapiens
; US-09-938-330-2

Query Match      29.3%; Score 48; DB 10; Length 451;
Best Local Similarity 45.0%; Pred. No. 70;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY      4  SHLGAGPAAACHGHLILGNP 23
        ||: ||||| |:: |
Db      54  SHVSGPAAASAGSMVVDTP 73

RESULT 10
US-09-938-330-6
; Sequence 6, Application US/09938330
; Patent No. US20020115838A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Hilbun, Erin
; APPLICANT: Scoville, John
; APPLICANT: Fiddle, Carl Johan
; APPLICANT: Hu, Yi
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. US20020115838A1 Human Proteases and Polynucleotides Enco
; FILE REFERENCE: LEX-0237-USA
; CURRENT APPLICATION NUMBER: US/09/938,330
; CURRENT FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 2001-08-22
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: US 60/227,104
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: US 60/233,796
; PRIOR FILING DATE: 2000-09-19
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 297
; TYPE: PRT
; ORGANISM: homo sapiens
; US-09-938-330-4
```

```
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: US 60/233,796
; PRIOR FILING DATE: 2000-09-19
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 486
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-938-330-6

Query Match      29.3%  Score 48; DB 10; Length 486;
Best Local Similarity 45.0%; Pred. No. 2.2e+02;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY      4 SHLGAGPAAACHGHLILGNP 23
      ||: ||||| |:: |
Db      54 SHVYSGPAAASAGSMVVDTP 73

RESULT 11
US-09-938-330-14
; Sequence 14, Application US/09938330
; Patent No. US20020115838A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Hilbun, Erin
; APPLICANT: Scoville, John
; APPLICANT: Friddle, Carl Johan
; APPLICANT: Hu, Yi
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. US20020115838A1el Human Proteases and Polynucleotides Encod
; FILE REFERENCE: LEX-0237-USA
; CURRENT APPLICATION NUMBER: US/09/938,330
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: US 60/227,104
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: US 60/233,796
; PRIOR FILING DATE: 2000-09-19
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 1213
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-938-330-14

Query Match      29.3%  Score 48; DB 10; Length 1213;
Best Local Similarity 45.0%; Pred. No. 2.2e+02;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY      4 SHLGAGPAAACHGHLILGNP 23
      ||: ||||| |:: |
Db      54 SHVYSGPAAASAGSMVVDTP 73

RESULT 12
US-09-938-330-12
; Sequence 12, Application US/09938330
; Patent No. US20020115838A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Hilbun, Erin
; APPLICANT: Scoville, John
; APPLICANT: Friddle, Carl Johan
; APPLICANT: Hu, Yi
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. US20020115838A1el Human Proteases and Polynucleotides Encod
; FILE REFERENCE: LEX-0237-USA
; CURRENT APPLICATION NUMBER: US/09/938,330
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: US 60/227,104
; PRIOR FILING DATE: 2000-08-22

; PRIOR APPLICATION NUMBER: US 60/233,796
; PRIOR FILING DATE: 2000-09-19
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 1213
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-938-330-12

Query Match      29.3%  Score 48; DB 10; Length 1213;
Best Local Similarity 45.0%; Pred. No. 2.2e+02;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY      4 SHLGAGPAAACHGHLILGNP 23
      ||: ||||| |:: |
Db      54 SHVYSGPAAASAGSMVVDTP 73

RESULT 13
US-09-938-330-10
; Sequence 10, Application US/09938330
; Patent No. US20020115838A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Hilbun, Erin
; APPLICANT: Scoville, John
; APPLICANT: Friddle, Carl Johan
; APPLICANT: Hu, Yi
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. US20020115838A1el Human Proteases and Polynucleotides Enco
; FILE REFERENCE: LEX-0237-USA
; CURRENT APPLICATION NUMBER: US/09/938,330
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: US 60/227,104
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: US 60/233,796
; PRIOR FILING DATE: 2000-09-19
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 1219
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-938-330-10

Query Match      29.3%  Score 48; DB 10; Length 1219;
Best Local Similarity 45.0%; Pred. No. 2.2e+02;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY      4 SHLGAGPAAACHGHLILGNP 23
      ||: ||||| |:: |
Db      54 SHVYSGPAAASAGSMVVDTP 73

RESULT 14
US-09-938-330-8
; Sequence 8, Application US/09938330
; Patent No. US20020115838A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Hilbun, Erin
; APPLICANT: Scoville, John
; APPLICANT: Friddle, Carl Johan
; APPLICANT: Hu, Yi
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. US20020115838A1el Human Proteases and Polynucleotides Enco
; FILE REFERENCE: LEX-0237-USA
; CURRENT APPLICATION NUMBER: US/09/938,330
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: US 60/227,104
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: US 60/233,796
; PRIOR FILING DATE: 2000-09-19
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 1219
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-938-330-8

Query Match      29.3%  Score 48; DB 10; Length 1219;
Best Local Similarity 45.0%; Pred. No. 2.2e+02;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY      4 SHLGAGPAAACHGHLILGNP 23
      ||: ||||| |:: |
Db      54 SHVYSGPAAASAGSMVVDTP 73
```


; PRIOR FILING DATE: 2000-09-19
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 1222
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-938-330-8

Query Match 29.3%; Score 48; DB 10; Length 1222;
Best Local Similarity 45.0%; Pred. No. 2.2e+02;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 4 SHLGAGPAAACHGHLILGNP 23
||: ||||| |:: |
DB 54 SHVSGPAAASAGSMVVDTP 73

RESULT 15

US-10-205-368-4
; Sequence 4, Application US/10205368
; Publication NO. US20030092621A1
; GENERAL INFORMATION:
; APPLICANT: AGOSTINO, MICHAEL J.
; APPLICANT: RACIE, LISA A.
; TITLE OF INVENTION: AGGREGANASE MOLECULES
; FILE REFERENCE: 08702.0094-00000
; CURRENT APPLICATION NUMBER: US/10/205,368
; CURRENT FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: 60/308,386
; PRIOR FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: 60/315,887
; PRIOR FILING DATE: 2001-08-29
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1223
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-205-368-4

Query Match 29.3%; Score 48; DB 9; Length 1223;
Best Local Similarity 45.0%; Pred. No. 2.2e+02;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 4 SHLGAGPAAACHGHLILGNP 23
||: ||||| |:: |
DB 54 SHVSGPAAASAGSMVVDTP 73

Search completed: June 3, 2003, 19:36:00
Job time : 8.43496 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2003, 19:20:46 ; Search time 3.40244 Seconds
(without alignments)
875.891 Million cell updates/sec

Title: US-09-938-703-6_COPY_185_215

Perfect score: 164

Sequence: 1 IKDHLGAGPAAACHGHLILGNPKNSASVSK 31

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_73:*

1: piri:*

2: piri:*

3: piri:*

4: piri:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Length | ID | Description |
|------------|-------|-------|--------|----------|--------------------|
| 1 | 56 | 34.1 | 240 | 2 S55480 | hypothetical prote |
| 2 | 55 | 33.5 | 336 | 2 AF2162 | UDP-glucose 4-epim |
| 3 | 53 | 32.3 | 499 | 2 D87363 | conserved hypothet |
| 4 | 52.5 | 32.0 | 478 | 2 T51274 | hypothetical prote |
| 5 | 52 | 31.7 | 408 | 2 H97236 | DNA repair exonuc |
| 6 | 49.5 | 30.2 | 597 | 2 J00107 | hypothetical 66K p |
| 7 | 49 | 29.9 | 331 | 2 T00582 | probable mitochond |
| 8 | 48.5 | 29.6 | 222 | 2 H82590 | conserved hypothet |
| 9 | 48.5 | 29.6 | 500 | 1 JE0259 | cytochrome P450 2D |
| 10 | 48 | 29.3 | 242 | 2 T32122 | hypothetical prote |
| 11 | 48 | 29.3 | 358 | 2 T08477 | inclusion membrane |
| 12 | 48 | 29.3 | 617 | 2 C72466 | probable translati |
| 13 | 47.5 | 29.0 | 660 | 2 G91021 | probable transform |
| 14 | 47.5 | 29.0 | 660 | 2 F85865 | probable transform |
| 15 | 47.5 | 29.0 | 660 | 2 E64996 | hypothetical prote |
| 16 | 47.5 | 29.0 | 2319 | 2 A47004 | coagulation factor |
| 17 | 47 | 28.7 | 294 | 2 G72706 | hypothetical prote |
| 18 | 47 | 28.7 | 302 | 2 T03109 | probable membrane |
| 19 | 47 | 28.7 | 830 | 2 A44768 | beta-glucosidase (|
| 20 | 46.5 | 28.4 | 516 | 2 S44191 | carboxypeptidase D |
| 21 | 46 | 28.0 | 388 | 2 T36766 | probable two-compo |
| 22 | 46 | 28.0 | 461 | 2 S23477 | probable benzoate |
| 23 | 46 | 28.0 | 469 | 2 T40339 | ferredoxin-NADP re |
| 24 | 46 | 28.0 | 478 | 2 T13683 | embryogenesis prot |
| 25 | 46 | 28.0 | 884 | 2 T18649 | hypothetical prote |
| 26 | 45.5 | 27.7 | 718 | 2 JC5805 | transcription fact |
| 27 | 45.5 | 27.7 | 745 | 2 B84673 | hypothetical prote |
| 28 | 45.5 | 27.7 | 816 | 2 F31277 | regulatory protein |
| 29 | 45.5 | 27.7 | 1733 | 1 B45344 | probable nuclear a |

30 45.5 27.7 1958 2 B40505
31 45.5 27.7 2351 1 EZHU
32 45 27.4 106 2 I48862
33 45 27.4 182 2 T06978
34 45 27.4 199 2 T44524
35 45 27.4 322 2 AE3093
36 45 27.4 329 2 D96030
37 45 27.4 340 2 E98193
38 45 27.4 376 2 A26066
39 45 27.4 390 2 T51197
40 45 27.4 437 2 A36372
41 45 27.4 484 2 T00158
42 45 27.4 511 2 H89758
43 45 27.4 638 2 T26490
44 45 27.4 675 2 I39065
45 45 27.4 815 2 S67675

ALIGNMENTS

RESULT 1

S55480

hypothetical protein SPAC5H10.02c - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C>Date: 01-Aug-1995 #sequence_revision 01-Sep-1995 #text_change 10-Dec-1999

C:Accession: S55480; T38966

R:Connor, R.; Churcher, C.M.

submitted to the EMBL Data Library, May 1995

A:Reference number: S55479

A:Accession: S55480

A>Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-240 <CON>

A:Cross-references: EMBL:Z49811; NID:g854599; PIDN:CAA89952.1; PID:g854601

R:Connor, R.; Churcher, C.M.; Barrrell, B.G.; Rajandream, M.A.; Walsby, S.V.

submitted to the EMBL Data Library, May 1995

A:Reference number: Z21821

A:Accession: T38966

A>Status: Preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-240 <CO2>

A:Cross-references: EMBL:Z49811; PIDN:CAA89952.1; GSPDB:GN00066; SPDB:SPAC5H10.02c

A:Experimental source: strain 972h-; cosmid c5H10

C:Genetics:

A:Gene: SPDB:SPAC5H10.02c

A:Map position: 1

C:Superfamily: conserved hypothetical protein YMR322c

Query Match 34.1%; Score 56; DB 2; Length 240;
Best Local Similarity 57.9%; Pred.No. 2.7;
Matches 11; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 11 AAACHGHLILGNPKNSASV 29

:||||:|||||

Db 138 SAVCHGPVLLANVKNPQSV 156

RESULT 2

AF2162

UDP-glucose 4-epimerase [Imported] - Nostoc sp. (strain PCC 7120)

C:Species: Nostoc sp.

A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C>Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002

C:Accession: AF2162

R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Irig

Nakazaki, N.; Shimp, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabat

DNA Res. 8, 205-213, 2001

A>Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium

A:Reference number: AB1807; MUID:21595285; PMID:11759840

A:Accession: AF2162

A>Status: Preliminary

A:Molecule type: DNA

```

RESULT 5
H97236
DNA repeat exonuclease [imported] - Clostridium acetobutylicum
C:Species: Clostridium acetobutylicum
C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 16-Aug-2002
C:Accession: H97236
R:Rolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; L.
; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 193, 4823-4838, 2001
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium
A:Reference number: A96900; MUID:21359325; PMID:21359325
A:Accession: H97236
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-408 <KUR>
A:Cross-references: GB:AF001437; PIDN:AAK80683.1; PID:g15025773; GSPDB:GN00168
A:Experimental source: Clostridium acetobutylicum ATCC824
C:Genetics:
C:Gene: CAC2737
C:Superfamily: exonuclease sbcd; phosphoesterase core homology

Query Match 31.7%; Score 52; DB 2; Length 408;
Best Local Similarity 45.0%; Pred. No. 16;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 6 LGAGPAAACHGHLILGNPKN 25
: ||| | ||::| ||:
DB 92 VVAGPLARDHGIIIVGTPKS 111

RESULT 6
JQ0107
hypothetical 66K protein - Ononis yellow mosaic virus
C:Species: Ononis yellow mosaic virus
A:Note: host Pisum sativum cv. Early massey (pea)
C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 08-Oct-1999
C:Accession: JQ0107
R:Ring, S.W.; Keese, P.; Gibbs, A.
Virology 172, 555-563, 1989
A:Title: Nucleotide sequence of the ononis yellow mosaic tymovirus genome.
A:Reference number: JQ0106; MUID:90021186; PMID:2800337
A:Accession: JQ0107
A:Molecule type: genomic RNA
A:Residues: 1-597 <DIN>
A:Cross-references: GB:J04375; NID:g332572; PIDN:AAA46795.1; PTD:g332573

Query Match 30.2%; Score 49.5; DB 2; Length 597;
Best Local Similarity 37.5%; Pred. No. 53;
Matches 12; Conservative 4; Mismatches 9; Indels 7; Gaps 1;

QY 4 SHLGAPAAACHGHLIL-----GNPKNSAS 28
||| : ||::| | ||::|
DB 553 SHLPSPSCSGDSFASCSFPGSPNPTSSAS 584

RESULT 7
T00582
probable mitochondrial carrier protein [imported] - Arabidopsis thaliana
N:Alternate names: hypothetical protein T2/E13.10
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 23-Mar-2001
C:Accession: T00582; C84705
R:Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; K.
submitted to the EMBL Data Library, May 1998
A:Description: Arabidopsis thaliana chromosome II BAC T2/E13 genomic sequence.
A:Reference number: Z14178
A:Accession: T00582
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-331 <ROD>
A:Cross-references: EMBL:AC004165; NID:g3150396; PID:g3150404
A:Experimental source: cultivar Columbia
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.

```

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanKen, S.E.; Umayam, L.; Tallon, L.; euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.

A:Reference number: R84420; MUID:20083487; PMID:10617197

A:Accession: C84705

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-331 <STO>

A:Cross-references: GB:AE002093; NID:g3150404; PIDN:AAAC16956.1; GSPDB:GN00139

C:Genetics:

A:Gene: At2g30160; T27E13.10

A:Map position: 2

A:Introns: 263/3

C:Superfamily: ADP.ATP carrier protein; ADP.ATP carrier protein repeat homology

C:Keywords: mitochondrion

Query Match 29.9%; Score 49; DB 2; Length 331;

Best Local Similarity 45.2%; Pred. No. 35;

Matches 14; Conservative 2; Mismatches 7; Indels 8; Gaps 1;

QY 6 LGAGPAAACH-----GHLLGNPKNSAS 28
||||| | | | | | | | | |

Db 104 LGAGPAHAYFSEVSKFLSGGNPNNSAA 134
||||| | | | | | | | | |

RESULT 8

H82590

conserved hypothetical protein XF2179 [imported] - Xylella fastidiosa (strain 9a5c)

C:Species: Xylella fastidiosa

C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 01-Mar-2002

C:Accession: H82590

R:Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing 406, 151-157, 2000

A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.

A:Reference number: A82515; MUID:20365717; PMID:10910347

A:Note: for a complete list of authors see reference number A59328 below

A:Accession: H82590

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-222 <SIM>

A:Cross-references: GB:AE004031; GB:AE003849; NID:g9107309; PIDN:AAF84978.1; GSPDB:GN001

A:Experimental source: Strain 9a5c

R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrier, H as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S. submitted to GenBank, June 2000

A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigh Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, B A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.; F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveira M.; Tsubako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z A:Reference number: A59328

A:Contents: annotation

C:Genetics:

A:Gene: XF2179

C:Superfamily: Bacillus subtilis conserved hypothetical protein yqeJ

Query Match 29.6%; Score 48.5; DB 2; Length 222;

Best Local Similarity 61.1%; Pred. No. 28;

Matches 11; Conservative 1; Mismatches 5; Indels 1; Gaps 1;

QY 15 HGHL-LLGPNKNSAVSK 31
||||| | | | | | | | | |

Db 170 HGHLWLLNQPPNSSASK 187
||||| | | | | | | | | |

RESULT 9

JE0259

cytochrome P450 2024 - rabbit

N:Contains: oxidoreductase (EC 1.-.-.-)

C:Species: Oryctolagus cuniculus (domestic rabbit)

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Mar-2000

C:Accession: JE0259

R:Yamamoto, Y.; Ishizuka, M.; Takada, A.; Fujita, S. J. Biochem. 124, 503-508, 1998

A:Title: Cloning, tissue distribution, and functional expression of two novel rabbit A:Reference number: JE0258; MUID:96391821; PMID:9722658

A:Accession: JE0259

A:Molecule type: mRNA

A:Residues: 1-500 <YAM>

A:Cross-references: DDBJ:AB008785

C:Comment: This protein shows high drug metabolizing activity.

C:Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology

C:Keywords: chromoprotein; electron transfer; endoplasmic reticulum; heme; iron; meti F:9-25/Domain: transmembrane #status predicted <TM1>

F:305-468/Domain: cytochrome P450 homology <CYP>

F:310-326/Domain: transmembrane #status predicted <TM2>

F:446/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 29.6%; Score 48.5; DB 1; Length 500;

Best Local Similarity 55.6%; Pred. No. 61;

Matches 10; Conservative 1; Mismatches 2; Indels 5; Gaps 1;

QY 5 HLGAGPAA-----ACHGH 17
||||| | | | | | | | | |

Db 112 HLGFQAGQVIMACYGH 129
||||| | | | | | | | | |

RESULT 10

T32122

hypothetical protein F59E11.5 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999

C:Accession: T32122

R:Bradshaw, H. submitted to the EMBL Data Library, July 1997

A:Description: The sequence of C. elegans cosmid F59E11.

A:Reference number: Z21124

A:Accession: T32122

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-242 <BRA>

A:Cross-references: EMBL:AF016685; PIDN:AA866219.1; GSPDB:GN00023; CESP:F59E11.5

A:Experimental source: strain Bristol N2; clone F59E11

C:Genetics:

A:Gene: CESP:F59E11.5

A:Map position: 5

A:Introns: 87/3; 123/3; 180/3; 217/3

Query Match 29.3%; Score 48; DB 2; Length 242;

Best Local Similarity 39.3%; Pred. No. 36;

Matches 11; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

QY 2 KDSHLGAGPAAACHGHLGNPKNSAV 29
||||| | | | | | | | | |

Db 88 KDELGRAAASLNHTQFLQRPHEVESL 115
||||| | | | | | | | | |

RESULT 11

T08477

Inclusion membrane protein C1 - Enterobacter aerogenes plasmid R751

C:Species: Enterobacter aerogenes

C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000

C:Accession: T08477

R:Thorsted, P.B.; Shah, D.S.; Macartney, D.; Kostelidou, K.; Thomas, C.M. Plasmid 36, 95-111, 1996

A:Title: Conservation of the genetic switch between replication and transfer genes o A:Reference number: Z16434; MUID:97118926; PMID:8954881

A:Accession: T08477

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-358 <THO>

Wed Jun 11 09:06:35 2003

A:Cross-references: EMBL:U67194; NID:gl572520; PIDN:ARC64421.1; PID:gl572526

C:Genetics:

A:Gene: incC1

A:Genome: plasmid R751

C:Superfamily: incC protein

Query Match 29.0%; Score 48; DB 2; Length 358;
 Best Local Similarity 58.8%; Pred. No. 52;
 Matches 10; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 IKDSHLGAGPAAACHGH 17

DB 14 IPDGHQAGDRAADHRH 30

RESULT 12

C72466 Probable translation initiation factor aIF-2 APE2374 [similarity] - Aeropyrum pernix (st

C:Species: Aeropyrum pernix

C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000

C:Accession: C72466

R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah

awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K

DNA Res. 6, 83-101, 1999

A>Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr

A:Reference number: A72450; MUID:99310339; PMID:10382966

A:Accession: C72466

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-617 <RAW>

A:Cross-references: DDBJ:AP000064; NID:g5105945; PIDN:BAA81387.1; PID:g5106076

A:Experimental source: strain K1

C:Genetics:

A:Gene: APE2374

C:Superfamily: translation initiation factor IF-2; translation elongation factor Tu hom

Query Match 29.3%; Score 48; DB 2; Length 617;

Best Local Similarity 27.6%; Pred. No. 88;

Matches 8; Conservative 8; Mismatches 13; Indels 0; Gaps 0;

QY 1 IKDSHLGAGPAAACHGHLLGNPKNSASV 29

DB 531 LEEARLGAAYVSIQGRILIGHRNEGDI 559

RESULT 13

G91021 Probable transformylase [imported] - Escherichia coli (strain O157:H7, substrain RMD 05

C:Species: Escherichia coli

C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001

C:Accession: G91021

R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.

gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.

DNA Res. 8, 11-22, 2001

A>Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene

A:Reference number: A99629; MUID:21156231; PMID:11258796

A:Accession: G91021

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-660 <RAY>

A:Cross-references: PIDN:BA836566.1; PID:gl3362613; GSPDB:GN00154

A:Experimental source: strain O157:H7, substrain RMD 050952

C:Genetics:

A:Gene: ECs3143

Query Match 29.0%; Score 47.5; DB 2; Length 660;

Best Local Similarity 42.1%; Pred. No. 1.1e+02;

Matches 8; Conservative 6; Mismatches 4; Indels 1; Gaps 1;

QY 14 CHGHLL-LGNPKNSASVSK 31

DB 558 CDGEIINIGNPENEASIEE 576

RESULT 14

F85865

Probable transformylase Z3513 [imported] - Escherichia coli (strain O157:H7, substrai

C:Species: Escherichia coli

C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001

C:Accession: F85865

R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May

iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apoda

Nature 409, 529-533, 2001

A>Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A:Reference number: A85480; MUID:21074935; PMID:11206551

A:Accession: F85865

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-660 <STO>

A:Cross-references: GB:AE005174; NID:gl2516601; PIDN:AA657386.1; GSPDB:GN00145; UWGP:

A:Experimental source: strain O157:H7, substrain EDL933

C:Genetics:

A:Gene: Z3513

Query Match 29.0%; Score 47.5; DB 2; Length 660;

Best Local Similarity 42.1%; Pred. No. 1.1e+02;

Matches 8; Conservative 6; Mismatches 4; Indels 1; Gaps 1;

QY 14 CHGHLL-LGNPKNSASVSK 31

DB 558 CDGEIINIGNPENEASIEE 576

RESULT 15

E64996

Hypothetical protein b2255 - Escherichia coli (strain K-12)

C:Species: Escherichia coli

C>Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 01-Mar-2002

C:Accession: E64996

R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;

A.; Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A>Title: The complete genome sequence of Escherichia coli K-12.

A:Reference number: A64720; MUID:97426617; PMID:9278503

A:Accession: E64996

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-660 <BLAT>

A:Cross-references: GB:AE000315; GB:U00096; NID:gl788582; PIDN:AACT5315.1; PID:gl7885

A:Experimental source: strain K-12, substrain MG1655

Query Match 29.0%; Score 47.5; DB 2; Length 660;

Best Local Similarity 42.1%; Pred. No. 1.1e+02;

Matches 8; Conservative 6; Mismatches 4; Indels 1; Gaps 1;

QY 14 CHGHLL-LGNPKNSASVSK 31

DB 558 CDGEIINIGNPENEASIEE 576

Search completed: June 3, 2003, 19:24:29

Job time : 5.40244 secs